Solutions Manual to Accompany:

Volume I: Fundamental Probability: A Computational Approach

Marc S. Paolella

Contents

Solutions to Chapter 1: Combinatorics

Solution to Problem 1.1:

a) There are 10 choices for the first place, and then 9 for the 2^{nd}, etc., so that there are $10! = 3.6288 \times 10^6$ possibilities.

b) Consider just the statistics books: there are $5! = 120$ ways of arranging them. Likewise, there are $3! = 6$ ways of arranging the econometrics books and $2! = 2$ ways for the German books. Thus, there are $120 \cdot 6 \cdot 2 = 1440$ ways of arranging the books in the order: statistics, econometrics, German. But, as the subject order is irrelevant, there are $1440 \cdot (3!) = 8640$ possible arrangements.

Solution to Problem 1.2:

a) 4!.

b) 1. The $\binom{4}{2,2} = 6$ possibilities are

#	A	B
1	$\{P_1, P_2\}$	$\{P_3, P_4\}$
2	$\{P_1, P_3\}$	$\{P_2, P_4\}$
3	$\{P_1, P_4\}$	$\{P_2, P_3\}$
4	$\{P_3, P_4\}$	$\{P_1, P_2\}$
5	$\{P_2, P_4\}$	$\{P_1, P_3\}$
6	$\{P_2, P_3\}$	$\{P_1, P_4\}$

2. That means the distinction between A and B is no longer there, so that there are only $6/2 = 3$ groups (either the first three or the last three in the above table.

Solution to Problem 1.3:

a) Of all the $11! = 3.9917 \times 10^7$ possible "people" combinations, no distinction can be made between the Germans, and likewise for the foreigners, so that there are only $11! \,/\, (8!\,3!) = 165$ combinations. (Why is $11! - 8! - 3!$ wrong?)

b) How many ways can one choose 5 objects from 11? Almost by definition, this is given by $\binom{11}{5} = \frac{11!}{5! \cdot (11-5)!} = 462$.

c) That means, we must pick 4 (from 8) Germans, and 1 (from 3) foreigners, or $\binom{8}{4}\binom{3}{1} = 70 \cdot 3 = 210$.

d) We can pick exactly 1 foreigner, which we just saw has 210 possibilities, or exactly 2 foreigners, for which there are $\binom{8}{3}\binom{3}{2} = 168$ ways, or we pick exactly 3 foreigners, with $\binom{8}{2}\binom{3}{3} = 28 \cdot 1 = 28$ ways, so that in total, we have $210 + 168 + 28 = 406$ ways. Alternatively, we could write $\binom{11}{5} - \binom{8}{5} = 406$. Similar to Example 1.2, observe that $\binom{3}{1}\binom{10}{4} = 630$ is not correct because duplication is not taken into account.

e) There are $\binom{3}{1} = 3$ ways of picking the foreigner, and we must choose 4 of the 8 Germans. If we pick neither of the feuding Germans, we have $\binom{6}{4} = 15$ possibilities. If we pick one of them, we have to choose 3 from $8 - 2$ Germans, or $\binom{6}{3}$, as well as one of the two (that is, $\binom{2}{1}$) feuders.

f) There are $\binom{3}{1}$ ways of picking the foreigner, so we must pick 4 Germans. Either we pick both " buddies", so we must choose 2 more from the remaining 6; or we pick neither of them, so we must pick 4 from the remaining 6. Thus we have

$$\binom{3}{1}\left\{\binom{2}{2}\binom{6}{2} + \binom{2}{0}\binom{6}{4}\right\} = 3\,(15 + 15) = 90$$

different ways.

Solution to Problem 1.4: (Note that (1.54) is precisely (1.12), just with a change of notation.) To prove (1.54) by induction over k, take $k = 0$ to get

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x},$$

showing that (1.54) is true for $k = 0$. Now assume the relation holds for k and we need to show that it holds for $k+1$, i.e., that

$$\sum_{n=0}^{\infty} \binom{n+k+1}{k+1} x^n = \frac{1}{(1-x)^{k+2}}$$

or, from (1.54) with k, that

$$\sum_{n=0}^{\infty} \binom{n+k+1}{k+1} x^n = \frac{1}{(1-x)} \frac{1}{(1-x)^{k+1}} = \frac{1}{(1-x)} \sum_{n=0}^{\infty} \binom{n+k}{k} x^n$$

or, with $h = n+1$, that

$$\begin{aligned}
(1-x)\sum_{n=0}^{\infty} \binom{n+k+1}{k+1} x^n &= \sum_{n=0}^{\infty} \binom{n+k+1}{k+1} x^n - \sum_{n=0}^{\infty} \binom{n+k+1}{k+1} x^{n+1} \\
&= \binom{k+1}{k+1} x^0 + \sum_{n=1}^{\infty} \binom{n+k+1}{k+1} x^n - \sum_{h=1}^{\infty} \binom{h+k}{k+1} x^h \\
&= 1 + \sum_{n=1}^{\infty} \left[\binom{n+k+1}{k+1} - \binom{n+k}{k+1}\right] x^n \\
&= 1 + \sum_{n=1}^{\infty} \binom{n+k}{k} x^n \\
&= \sum_{n=0}^{\infty} \binom{n+k}{k} x^n \\
&= \frac{1}{(1-x)^{k+1}}
\end{aligned}$$

from (1.4) and (1.54) with k. Dividing both sides by $1 - x$ finishes the induction argument.

Solution to Problem 1.5: From (1.11) with $t = -1/2$ and $x = 4a$,

$$\frac{1}{\sqrt{1-4a}} = (1-4a)^{-1/2} = \sum_{i=0}^{\infty} \binom{-1/2}{i} (-4a)^i$$

and, from (1.10),

$$\sum_{i=0}^{\infty} \binom{-1/2}{i} (-4a)^i = \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^{2i} (-1)^i \binom{2i}{i} (-4a)^i = \sum_{i=0}^{\infty} \binom{2i}{i} a^i.$$

Solution to Problem 1.6: From (1.4), showing

$$\sum_{j=i}^{l} (-1)^j \binom{l}{j} = (-1)^i \binom{l-1}{i-1}$$

is equivalent to showing that

$$\sum_{j=i}^{l} (-1)^j \binom{l-1}{j-1} + \sum_{j=i}^{l} (-1)^j \binom{l-1}{j} - (-1)^i \binom{l-1}{i-1} = 0$$

or

$$\sum_{j=i}^{l} (-1)^j \binom{l-1}{j-1} + \sum_{j=i}^{l} (-1)^j \binom{l-1}{j} + (-1)^{i-1} \binom{l-1}{i-1} = 0$$

or

$$\sum_{j=i}^{l} (-1)^j \binom{l-1}{j-1} + \sum_{j=i-1}^{l} (-1)^j \binom{l-1}{j} = 0$$

or, substituting $k = j+1$ into the second term and factoring out a -1,

$$\sum_{j=i}^{l} (-1)^j \binom{l-1}{j-1} + (-1) \sum_{k=i}^{l+1} (-1)^k \binom{l-1}{k-1} = 0.$$

Replacing k with j and recalling that, in general, $\binom{n}{k} = 0$ for $n < k$, the last term in the second sum is zero, so that we get

$$\sum_{j=i}^{l} (-1)^j \binom{l-1}{j-1} + (-1) \sum_{j=i}^{l} (-1)^j \binom{l-1}{j-1},$$

which is identically zero, as was to be shown.

Solution to Problem 1.7:

a) From the binomial theorem (1.18),

$$(1-p)^{n-j} = \sum_{k=0}^{n-j} \binom{n-j}{k} (-1)^k p^k$$

so that

$$B_{i,n} = \sum_{j=i}^{n} \sum_{k=0}^{n-j} (-1)^k \binom{n-j}{k} \binom{n}{j} p^{j+k}.$$

From the following tables of the indices,

j	k				
i	0	1	2	$\cdots$	$n-i$
$i+1$	0	1	$\cdots$	$n-i-1$	
$\vdots$					
$n-1$		/			
n	0				

j	$l = j+k$				
i	i	$i+1$	$i+2$	$\cdots$	n
$i+1$	$i+1$	$i+2$	$\cdots$	n	
$\vdots$					
$n-1$	$n-1$	/			
n	n				

it is clear that, for $l = j+k$,

$$B_{i,n} = \sum_{l=i}^{n} \left\{ \sum_{j=i}^{l} (-1)^{l-j} \binom{n-j}{l-j} \binom{n}{j} \right\} p^l.$$

Then $A_{i,n}$ and $B_{i,n}$ are equivalent if the coefficients of like powers of p are the same in both expressions, i.e., if

$$(-1)^{l-i}\binom{l-1}{i-1}\binom{n}{l} = \sum_{j=i}^{l}(-1)^{l-j}\binom{n-j}{l-j}\binom{n}{j}$$

or, as

$$\binom{n-j}{l-j}\binom{n}{j} = \frac{(n-j)!}{(l-j)!\,(n-l)!}\frac{n!}{j!\,(n-j)!}\frac{l!}{l!} = \frac{l!}{(l-j)!j!}\frac{n!}{l!\,(n-l)!} = \binom{l}{j}\binom{n}{l},$$

if

$$\sum_{j=i}^{l}(-1)^{j}\binom{l}{j} = (-1)^{i}\binom{l-1}{i-1}.$$

This was, however, shown in Problem 1.6, so that $A_{i,n}$ and $B_{i,n}$ are indeed equal.

b) To show that $A_{i,n} = B_{i,n}$ for all $n \in \mathbb{N}$ and $1 \le i \le n$, use induction on $n-i$. For $n-i=0$, i.e., $i=n$, the assertion is true; one easily checks

$$A_{n,n} = p^n = B_{n,n}. \tag{S-1.1}$$

It suffices to show $A_{i,n} - A_{i+1,n} = B_{i,n} - B_{i+1,n}$ for all $1 \le i < n$ because then $A_{i,n} - B_{i,n} = A_{i+1,n} - B_{i+1,n}$ and, from (S-1.1), $A_{n-1,n} - B_{n-1,n} = 0$, etc. For $1 \le n < n$,

$$\begin{aligned}
A_{i,n} - A_{i+1,n} &= \sum_{j=i+1}^{n}(-1)^{j-i}\binom{n}{j}p^j\left(\binom{j-1}{i-1}+\binom{j-1}{i}\right)+\binom{n}{i}p^i \\
&= \sum_{j=i+1}^{n}(-1)^{j-i}\binom{n}{j}p^j\binom{j}{i}+\binom{n}{i}p^i = \sum_{j=i}^{n}(-1)^{j-i}\binom{n}{j}p^j\binom{j}{i} \\
&= \sum_{k=0}^{n-i}(-1)^k\binom{n}{i+k}p^{i+k}\binom{i+k}{i} = p^i\sum_{k=0}^{n-i}(-1)^k\binom{n}{i}p^k\binom{n-i}{k} \\
&= \binom{n}{i}p^i\sum_{k=0}^{n-i}\binom{n-i}{k}(-p)^k = \binom{n}{i}p^i(1-p)^{n-i},
\end{aligned}$$

using (1.4) and (1.18). On the other hand, it is easy to see that

$$B_{i,n} - B_{i+1,n} = \binom{n}{i}p^i(1-p)^{n-i},$$

and we are done.

Solution to Problem 1.8: Substitute $i = y-k$ so that

$$\begin{aligned}
\sum_{i=0}^{y}\binom{2i}{y}\binom{y}{i}(-1)^{i+y} &= \sum_{k=y}^{0}\binom{2y-2k}{y}\binom{y}{y-k}(-1)^{-k} \\
&= \sum_{k=0}^{y}\binom{y-k}{k}\binom{2y-2k}{y-k}(-1)^{-k},
\end{aligned}$$

where the latter equality follows because

$$\begin{aligned}
\binom{2y-2k}{y}\binom{y}{y-k} &= \frac{(2y-2k)!}{y!\,(y-2k)!}\frac{y!}{(y-k)!k!} \\
&= \frac{(y-k)!}{k!\,(y-2k)!}\frac{(2y-2k)!}{(y-k)!\,(y-k)!} = \binom{y-k}{k}\binom{2y-2k}{y-k}.
\end{aligned}$$

Either binomial coefficient $\binom{2y-2k}{y}$ or $\binom{y-k}{k}$ implies $k \le y/2$ which, for y even or odd, is given by $k \le \lfloor y/2 \rfloor$.

Solution to Problem 1.9: Expressing the factorials as a binomial coefficient and letting $x = y - 1$,

$$\sum_{y=1}^{n} \frac{(n-1)!\,y}{(n-y)!\,n^y} = \sum_{y=1}^{n} \binom{n-1}{y-1} \frac{y!}{n^y}$$
$$= \sum_{x=0}^{n-1} \binom{n-1}{x} \frac{(x+1)!}{n^{x+1}} = \frac{1}{n} \sum_{x=0}^{n-1} \binom{n-1}{x} \frac{(x+1)!}{n^x}.$$

Next, with $i = n - 1 - x$, this is

$$\frac{1}{n} \sum_{i=n-1}^{0} \binom{n-1}{n-1-i} \frac{(n-i)!}{n^{n-1-i}} = \frac{1}{n^n} \sum_{i=0}^{n-1} \binom{n-1}{i} (n-i)! n^i$$
$$= \frac{1}{n^n} \sum_{i=0}^{n-1} \frac{(n-1)!}{i!} (n-i)\, n^i. \tag{S-1.2}$$

Splitting up the $n - i$ term gives

$$\frac{1}{n^n} \left\{ \sum_{i=0}^{n-1} \frac{n!}{i!} n^i - \sum_{i=1}^{n-1} \frac{(n-1)!}{(i-1)!} n^i \right\} = \frac{(n-1)!}{n^n} \left\{ \sum_{i=0}^{n-1} \frac{n^{i+1}}{i!} - \sum_{i=1}^{n-1} \frac{n^i}{(i-1)!} \right\}$$

or, with $j = i - 1$,

$$\frac{(n-1)!}{n^n} \left\{ \sum_{i=0}^{n-1} \frac{n^{i+1}}{i!} - \sum_{j=0}^{n-2} \frac{n^{j+1}}{j!} \right\} = \frac{(n-1)!}{n^n} \left\{ \frac{n^n}{(n-1)!} \right\} = 1.$$

Knowing that the sum is one, it follows directly from (S-1.2) that

$$n^n = \sum_{i=0}^{n-1} \binom{n-1}{i} (n-i)! n^i.$$

Solution to Problem 1.10:

a) From (1.5),

$$\binom{r+m+1}{m} = \sum_{j=0}^{m} \binom{(r+m+1)-j-1}{m-j} = \sum_{i=0}^{m} \binom{r+i}{i}$$

using the substitution $i = m - j$ and reversing the order of summation.

b) From (1.28),

$$\binom{n+m}{k} = \sum_{i=0}^{k} \binom{n}{i} \binom{m}{k-i}, \quad k \le n,\ k \le m$$

and with $m = n$ and $k = n$ and the fact that $\binom{n}{n-i} = \binom{n}{i}$, the result follows.

c) For $M = 1$,

$$\sum_{i=0}^{1} \binom{N}{k+i} \binom{1}{i} = \binom{N}{k} + \binom{N}{k+1} = \binom{N+1}{k+1}$$

from (1.4). Assume that the identity holds for M. Then, for $M + 1$,

$$\sum_{i=0}^{M+1} \binom{N}{k+i} \binom{M+1}{i}$$

is equal to

$$\begin{aligned}
&\binom{N}{k+M+1} + \sum_{i=0}^{M} \binom{N}{k+i} \left\{ \binom{M}{i} + \binom{M}{i-1} \right\} \\
&= \binom{N}{k+M+1} + \sum_{i=0}^{M} \binom{N}{k+i} \binom{M}{i} + \sum_{i=0}^{M} \binom{N}{k+i} \binom{M}{i-1} \\
&= \binom{N}{k+M+1} + \binom{N+M}{k+M} + \sum_{i=0}^{M-1} \binom{N}{k+i+1} \binom{M}{i} \\
&= \binom{N}{k+M+1} + \binom{N+M}{k+M} + \binom{N+M}{k+1+M} - \binom{N}{k+M+1} \\
&= \binom{N+M+1}{k+M+1},
\end{aligned}$$

again from (1.4) and the proof is complete. The second method of proof is as follows. From the binomial theorem, the lhs of (1.56) is

$$\sum_{i=0}^{N} \sum_{j=0}^{M} \binom{N}{i} \binom{M}{j} x^{i-j}$$

with coefficient of x^k being

$$\sum_{i=0}^{N} \binom{N}{i} \binom{M}{i-k} = \sum_{j=0}^{M} \binom{N}{k+j} \binom{M}{j}.$$

The result now follows because the rhs of (1.56) is

$$\sum_{i=0}^{N+M} \binom{N+M}{i} x^{i-M},$$

with the coefficient of x^k being $\binom{N+M}{k+M}$.

d) Let

$$A = \sum_{j=0}^{r} \binom{r}{j} a^{r-j} (b-a)^j \frac{1}{j+1}.$$

Multiply A by $r+1$ and $b-a$ and use

$$\binom{r}{j} \frac{r+1}{j+1} = \frac{(r+1)!}{(r-j)!\,(j+1)!} = \binom{r+1}{j+1}$$

to get, with $q = j+1$ and $s = r+1$,

$$\begin{aligned}
(b-a)(r+1)A &= \sum_{j=0}^{r} \binom{r+1}{j+1} a^{r-j} (b-a)^{j+1} \\
&= \sum_{q=1}^{s} \binom{s}{q} a^{s-q} (b-a)^q \\
&= \sum_{q=0}^{s} \binom{s}{q} a^{s-q} (b-a)^q - \binom{s}{0} a^{s-0} (b-a)^0 \\
&= b^s - a^s = b^{r+1} - a^{r+1},
\end{aligned}$$

i.e.,

$$A = \frac{b^{r+1} - a^{r+1}}{(b-a)(r+1)}.$$

Solution to Problem 1.11:

a) From (1.9), it follows that

$$\binom{r_1+i-1}{i} = (-1)^i \binom{-r_1}{i} \quad \text{and} \quad \binom{r_2+y-i-1}{y-i} = (-1)^{y-i} \binom{-r_2}{y-i},$$

so that the rhs of the desired equation is

$$\begin{aligned} S &= \sum_{i=0}^{y} \binom{-r_1}{i} \binom{-r_2}{y-i} (-1)^y \\ &\overset{(1.28)}{=} (-1)^y \binom{-(r_1+r_2)}{y} \overset{(1.9)}{=} \binom{r_1+r_2+y-1}{y}. \end{aligned}$$

b) The desired equality is equivalent to stating

$$2^{2N} = \sum_{k=0}^{N} \binom{2N-k}{N} 2^k = \sum_{i=0}^{N} \binom{N+i}{N} 2^{N-i} = 2^N \sum_{i=0}^{N} \binom{N+i}{N} \left(\frac{1}{2}\right)^i,$$

or

$$2^N = \sum_{i=0}^{N} \binom{N+i}{N} \left(\frac{1}{2}\right)^i. \tag{S-1.3}$$

But (S-1.3) holds, because, from (1.14), it follows that

$$2^{n-1} = \sum_{i=0}^{n-1} \binom{n+i-1}{i} \left(\frac{1}{2}\right)^i,$$

which is the same as (S-1.3) with $N = n-1$. A probabilistic proof of this is given in Example 4.13.

c) From (1.4),

$$\begin{aligned} \sum_{i=0}^{n-1} \binom{2n}{i} &= \sum_{i=0}^{n-1} \left\{ \binom{2n-1}{i} + \binom{2n-1}{i-1} \right\} = \sum_{i=0}^{n-1} \binom{2n-1}{i} + \sum_{i=0}^{n-2} \binom{2n-1}{i} \\ &= \binom{2n-1}{n-1} + 2 \sum_{i=0}^{n-2} \binom{2n-1}{i}, \end{aligned}$$

and applying this recursively gives

$$\begin{aligned} \sum_{i=0}^{n-1} \binom{2n}{i} &= \binom{2n-1}{n-1} + 2 \sum_{i=0}^{n-2} \binom{2n-1}{i} \\ &= \binom{2n-1}{n-1} + 2 \left\{ \binom{2n-2}{n-2} + 2 \sum_{i=0}^{n-3} \binom{2n-2}{i} \right\} \\ &\vdots \\ &= \binom{2n-1}{n-1} + 2\binom{2n-2}{n-2} + 4\binom{2n-3}{n-3} + \cdots + \binom{n}{0} \\ &= \sum_{i=0}^{n-1} 2^i \binom{2n-1-i}{n-1-i}. \end{aligned}$$

Solution to Problem 1.12:

a) For $n=0$,

$$\begin{aligned}\sum_{i=1}^{N-1}\binom{N}{i}(-1)^{i+1} &= (-1)\sum_{i=0}^{N}\binom{N}{i}(-1)^{i}-\binom{N}{0}(-1)^{0+1}-\binom{N}{N}(-1)^{N+1}\\ &= (-1)\cdot 0+1+(-1)^{N}\\ &= \begin{cases}2, & \text{if } N \text{ is even,}\\ 0, & \text{if } N \text{ is odd.}\end{cases}\end{aligned}$$

from the binomial theorem (with $x=-1$ and $y=1$). As an aside, note that (1.60) can equivalently be expressed as

$$0=\sum_{i=0}^{N}\binom{N}{i}\left(\frac{N-i}{N}\right)^{n}(-1)^{i+1}, \quad 1\le n\le N-1, \tag{S-1.4}$$

but that, for $n=0$, one has to be careful how 0^0 is treated. See the remark on page 36.

For $n=1$, with

$$\binom{N}{i}\left(\frac{N-i}{N}\right)=\binom{N-1}{i},$$

it follows from the binomial theorem that

$$\sum_{i=1}^{N-1}\binom{N}{i}\left(\frac{N-i}{N}\right)(-1)^{i+1}=1-\sum_{i=0}^{N-1}\binom{N-1}{i}(-1)^{i}=1.$$

Before starting the induction proof, we calculate the sum in (1.60) for $N=n$. For values $N=2,3,4,5,10$ and 20, Maple gives $\frac{1}{2}$, $\frac{7}{9}$, $\frac{29}{32}$, $\frac{601}{625}$, $\frac{1561933}{1562500}$ and $\frac{639999985150744579}{640000000000000000}\approx 1-2.3202\times10^{-8}$, respectively, so that the identity does not hold for $N=n$ (or, in fact, for any $n>N$), but quickly approaches one.

For the induction argument, $P(n,N)$ is equal to

$$\begin{aligned}&\sum_{i=1}^{N-1}\binom{N}{i}\left(\frac{N-i}{N}\right)^{n}(-1)^{i+1}\\ &=\sum_{i=1}^{N-1}\binom{N-1}{i}\left(\frac{N-i}{N}\right)^{n}(-1)^{i+1}+\sum_{i=1}^{N-1}\binom{N-1}{i-1}\left(\frac{N-i}{N}\right)^{n}(-1)^{i+1}\\ &=P(n+1,N)+Z,\end{aligned} \tag{S-1.5}$$

where the first term follows by writing

$$\left(\frac{N-i}{N}\right)^{n}=\left(\frac{N}{N-i}\right)\left(\frac{N-i}{N}\right)^{n+1}$$

and combining into the combinatoric, and the second term in (S-1.5) is labeled Z just for convenience. Thus, in order for (1.60) to hold, Z must be zero. But, with $j=i-1$,

$$\begin{aligned}Z &= \sum_{i=1}^{N-1}\binom{N-1}{i-1}\left(\frac{N-i}{N}\right)^{n}(-1)^{i+1}=\sum_{j=0}^{N-2}\binom{N-1}{j}\left(\frac{N-j-1}{N}\right)^{n}(-1)^{j}\\ &= \left(\frac{N-1}{N}\right)^{n}\sum_{j=0}^{N-2}\binom{N-1}{j}\left(\frac{N-j-1}{N-1}\right)^{n}(-1)^{j}\\ &= \left(\frac{N-1}{N}\right)^{n}\left(1+(-1)\sum_{j=1}^{N-2}\binom{N-1}{j}\left(\frac{N-j-1}{N-1}\right)^{n}(-1)^{j+1}\right)\\ &= \left(\frac{N-1}{N}\right)^{n}(1-P(n,N-1)).\end{aligned}$$

This last expression is zero if $P(n, N-1) = 1$, which holds only for $n \leq N-2$. In that case, $P(n, N) = P(n+1, N)$, which is one from the result for $P(1, N)$, but only if $n+1 \leq N-1$ or, again, $n \leq N-2$. The fact that the result also holds for $n = N-1$ is not in this proof and would have to be shown separately. For example, with $N = 5$ and $n = 4$, $P(n, N)$ is indeed 1, but $P(n+1, N) = 601/625$ and $Z = 24/625$.

An alternative proof which encompasses the $n = N-1$ case is as follows. Directly from the above decomposition, we have

$$P(n+1, N) = P(n, N) - \left(\frac{N-1}{N}\right)^n (1 - P(n, N-1)),$$

so that, changing $n+1$ to n,

$$P(n, N) = P(n-1, N) - \left(\frac{N-1}{N}\right)^{n-1} (1 - P(n-1, N-1)). \qquad \text{(S-1.6)}$$

From this, $P(n, N)$ can be recursively calculated using the initial conditions $P(1, a) = 1$ for $a \geq 2$. So, assuming $P(n-1, N) = 1$ for $n-1 \leq (N-1)-1$, then, from (S-1.6),

$$P(n, N) = 1 - \left(\frac{N-1}{N}\right)^{n-1} (1-1) = 1$$

for $n \leq N-1$. (Notice that, if $n \leq N-1$, then $n-1 \leq N-1$ and $n-1 \leq (N-1)-1$.)

b) Starting from (S-1.4), this can be easily expressed as

$$0 = N^{-n} \sum_{i=0}^{N} \binom{N}{N-i} (N-i)^n (-1)^{i+1} = \sum_{j=0}^{N} \binom{N}{j} j^n (-1)^{N-j+1},$$

for $1 \leq n \leq N-1$, with $j = N-i$ and reversing the summation order. The next exercise shows that $Q(N, N) = -N!$, which differs greatly from 0 as N grows!

c) With $0^0 = 1$, $M(0) = 1$ and it is simple to see that $M(1) = 1$. Then, for $N+1$,

$$\begin{aligned} M(N+1) &= \sum_{j=1}^{N+1} \binom{N+1}{j} j^{N+1} (-1)^{N-j+1} \\ &= \sum_{i=0}^{N} \binom{N+1}{i+1} (i+1)^{N+1} (-1)^{N-i} \\ &= (N+1) \sum_{i=0}^{N} \binom{N}{i} (i+1)^N (-1)^{N-i} \\ &= (N+1) M(N) = (N+1)!. \end{aligned}$$

Solution to Problem 1.13: To prove (1.61), i.e., the identity

$$\sum_{i=0}^{y} \binom{n+y}{y-i} (-1)^i = \binom{n+y-1}{y}$$

for $n \geq 1$ and $y \geq 0$, first note that it holds for $y = 0$ and $y = 1$. To use induction, assume (1.61) holds for y, define $h_y = \binom{n+y-1}{y}$ for convenience and use (1.4) to get

$$\sum_{i=0}^{y+1} \binom{n+(y+1)}{(y+1)-i} (-1)^i = \sum_{i=0}^{y+1} \left\{ \binom{n+y}{y-i} + \binom{n+y}{y+1-i} \right\} (-1)^i.$$

Then, using (1.61) (which is allowed, because we *assume* it holds for y), the previous expression is, with $j = i - 1$,

$$
\begin{aligned}
& h_y + \sum_{i=0}^{y+1} \binom{n+y}{y+1-i} (-1)^i = h_y + \binom{n+y}{y+1} + \sum_{i=1}^{y+1} \binom{n+y}{y+1-i} (-1)^i \\
= \;& h_y + \binom{n+y}{y+1} + \sum_{j=0}^{y} \binom{n+y}{y-j} (-1)^{j+1} \\
= \;& h_y + \binom{n+y}{y+1} + (-1) \sum_{j=0}^{y} \binom{n+y}{y-j} (-1)^{j} \\
= \;& h_y + h_{y+1} - h_y = \binom{n+y}{y+1}.
\end{aligned}
$$

Solution to Problem 1.14:

a) Using (1.6) expressed as

$$\sum_{i=0}^{k} \binom{r+i}{i} = \binom{r+k+1}{k},$$

the rhs of (1.17) for $i = 1$ is

$$
\begin{aligned}
& (N-n) \binom{N}{n} - \sum_{x=1}^{N-n} x \binom{n+x-1}{x} \\
= \;& N \binom{N}{n} - n \left\{ \binom{N}{n} + \binom{N}{n+1} \right\} \\
= \;& N \binom{N}{n} - n \binom{N+1}{n+1} = \binom{N}{n} \left\{ N - \frac{n\,(N+1)}{n+1} \right\} \\
= \;& \binom{N}{n} \frac{N-n}{n+1} = \binom{N}{n+1},
\end{aligned}
$$

because

$$
\begin{aligned}
\sum_{x=1}^{N-n} x \binom{n+x-1}{x} &= \sum_{x=1}^{N-n} \frac{(n+x-1)\cdots n}{(x-1)!} = \sum_{x=0}^{N-n-1} \frac{(n+x)\cdots n}{x!} \\
&= n \sum_{x=0}^{N-n-1} \binom{n+x}{x} = n \binom{N}{N-n-1} = n \binom{N}{n+1}.
\end{aligned}
$$

Now define $Z\,(i)$ to be the rhs of (1.17) and assume that (1.17) is valid for $i - 1$. Then, with $\binom{a}{b} = 0$ for $a < b$,

$$
\begin{aligned}
Z\,(i) &= \frac{1}{i} \sum_{x=0}^{N-n} \binom{n+x-1}{x} \binom{N-n-x}{i-1} (N-n-x-i+1) \\
&= \frac{(N-n)-(i-1)}{i} Z\,(i-1) - \frac{1}{i} \sum_{x=0}^{N-n} x \binom{n+x-1}{x} \binom{N-n-x}{i-1}
\end{aligned}
$$

and

$$
\begin{aligned}
& \frac{1}{i} \sum_{x=0}^{N-n} x \binom{n+x-1}{x} \binom{N-n-x}{i-1} = \frac{n}{i} \sum_{x=1}^{N-n} \binom{n+x-1}{x-1} \binom{N-n-x}{i-1} \\
= \;& \frac{n}{i} \sum_{x=0}^{N-(n+1)} \binom{n+x}{x} \binom{N-(n+1)-x}{i-1} = \frac{n}{i} \binom{N}{(n+1)+(i-1)}
\end{aligned}
$$

so that

$$
\begin{aligned}
Z(i) &= \frac{N-n-i+1}{i}\binom{N}{n+i-1} - \frac{n}{i}\binom{N}{n+i} \\
&= \left\{\frac{(N-n-i+1)(n+i)}{(N-n-i+1)\cdot i} - \frac{n}{i}\right\}\binom{N}{n+i} = \binom{N}{n+i},
\end{aligned}
$$

proving (1.17).

b) For the lhs of (1.62),

$$
\begin{aligned}
&\sum_{i=0}^{y}\binom{N}{n+i}\binom{N-n-i}{y-i}(-1)^i \\
=\ &\binom{N}{n}\binom{N-n}{y} - \binom{N}{n+1}\binom{N-n-1}{y-1} + \cdots + (-1)^y\binom{N}{n+y} \\
=\ &\binom{N}{n+y}\left\{\binom{n+y}{y} - \binom{n+y}{y-1} + \binom{n+y}{y-2} - \cdots + (-1)^y\right\} \\
=\ &\binom{N}{n+y}\sum_{i=0}^{y}\binom{n+y}{y-i}(-1)^i = \binom{N}{n+y}\binom{n+y-1}{y},
\end{aligned}
$$

from (1.61). For the rhs of (1.62),

$$
\begin{aligned}
&\sum_{i=y}^{N-n}\binom{n+i-1}{i}\binom{i}{y} = \frac{1}{y!}\sum_{i=y}^{N-n}\frac{(n+i-1)(n+i-2)\cdots n}{(i-y)!} \\
=\ &\frac{n(n+1)\cdots(n+y-1)}{y!}\sum_{i=y}^{N-n}\frac{(n+i-1)\cdots(n+y)}{(i-y)!} \\
=\ &\binom{n+y-1}{y}\sum_{j=0}^{N-n-y}\binom{n+j+y-1}{j} = \binom{n+y-1}{y}\binom{N}{n+y},
\end{aligned}
$$

where the second to last summand is one for $i=y$, $(n+y)$ for $i=y+1$, etc.

c) Cancel p^n from both sides of (1.63) and substitute

$$
(1-p)^i = \sum_{k=0}^{i}\binom{i}{k}(-1)^k p^k
$$

into the lhs of (1.63) and

$$
(1-p)^{N-n-i} = \sum_{k=0}^{N-n-i}\binom{N-n-i}{k}(-1)^k p^k
$$

into the rhs, to get

$$
\sum_{i=0}^{N-n}\sum_{k=0}^{i}\binom{n+i-1}{i}\binom{i}{k}(-1)^k p^k \overset{?}{=} \sum_{i=0}^{N-n}\sum_{k=0}^{N-n-i}\binom{N}{n+i}\binom{N-n-i}{k}(-1)^k p^{i+k}, \qquad \text{(S-1.7)}
$$

with equality holding when the coefficients of p^c coincide in both polynomials, $c = 0, 1, \ldots, N-n$. The double sum on the lhs of (S-1.7) is the same as $\sum_{k=0}^{N-n}\sum_{i=k}^{N-n}$, so that, substituting $y=k$, the lhs of (S-1.7) is

$$
\sum_{y=0}^{N-n}(-1)^y\sum_{i=y}^{N-n}\binom{n+i-1}{i}\binom{i}{y}p^y. \qquad \text{(S-1.8)}
$$

The order change in the double sum is easily seen by using the table below. The double sum $\sum_{i=0}^{N-n}\sum_{k=0}^{i}$ starts with the leftmost column and works downwards, while $\sum_{k=0}^{N-n}\sum_{i=k}^{N-n}$ starts with the top row and counts across.

	i: 0	1	2	$\cdots$	$N-n$
	0	0	0		0
		1	1		1
k			2		2
				$\searrow$	$\vdots$
					$N-n$

Now consider the rhs of (S-1.7), which by similar reasoning is equal to

$$\sum_{y=0}^{N-n}(-1)^y\sum_{i=0}^{y}\binom{N}{n+i}\binom{N-n-i}{y-i}(-1)^i p^y. \qquad \text{(S-1.9)}$$

The result now follows, because the equality of (S-1.8) and (S-1.9) is the same as (1.62).

A probabilistic proof of (1.63) is given in §3.3.

Solution to Problem 1.15:

a) Listing S-1.1 gives a program which implements the switching method. The following code was run to simulate the performance.

```
N=30; s=1000; m = zeros(s,N);
for i=1:s, m(i,:)=permvecswitch(N); end
```

The subsequent command `plot(1:30,mean(store),'r-o')` produced Figure S-1.1, which shows the mean (over the 1000 simulations) of each column of matrix `store`. We see clearly that there is a tendency for the initial elements of the random permutation vector to be small. Figure S-1.2 shows the frequencies associated with the first element of the output vector. It was produced with `xx=tabulate(store(:,1)); plot(xx(:,1),xx(:,2),'ro'); grid`, where `tabulate` is a built–in Matlab function which computes the frequencies of elements in a vector.

```
function y = permvecswitch(N)
  y=1:N;
  for i=1:N
  p = unidrnd(N);
  temp = y(i); y(i) = y(p); y(p)=temp; % switch the two elements
end
```

Program Listing S-1.1: Returns a random permutation of vector $(1, 2, \ldots, N)$ by switching N pairs of elements

b) Listing S-1.2 shows one method of accomplishing this. It simply switches each value which overlaps with a another randomly chosen element of the vector.

Solution to Problem 1.16: With $u = \dfrac{1}{1+x}$, $x = \dfrac{1-u}{u}$ and $dx = -\dfrac{1}{u^2}\,du$,

$$\begin{aligned} I &= -\int_1^0 \frac{1}{\left(\frac{1-u}{u}\right)^{1/2}\left(\frac{1}{u}\right)}\frac{1}{u^2}du \\ &= \int_0^1 u^{-1/2}(1-u)^{-1/2}\,du = B\left(\frac{1}{2},\frac{1}{2}\right) = \frac{\Gamma(1/2)\,\Gamma(1/2)}{\Gamma(1)} = \pi. \end{aligned}$$

This integral (and many others) can also be resolved using contour integration (see, e.g., Bak and Newman, 1997, p. 138).

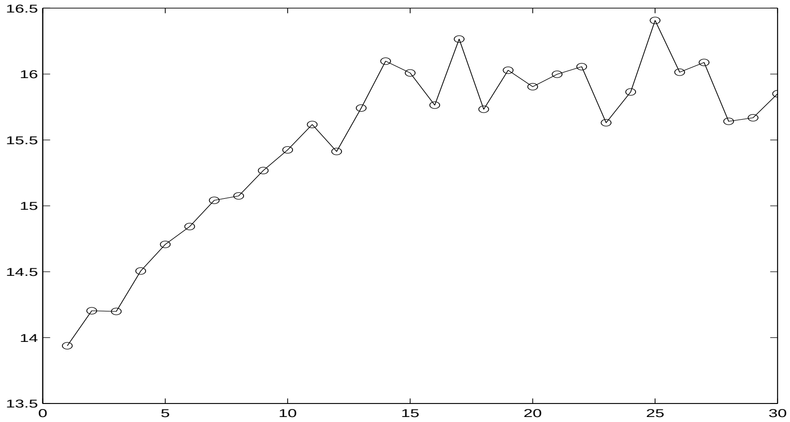


Figure S-1.1: Mean for each of the 30 positions computed from 1000 simulated output vectors of `permvecswitch(30)`

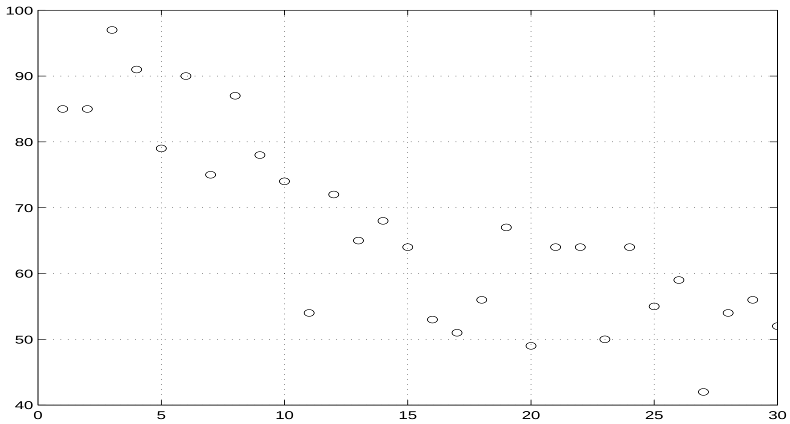


Figure S-1.2: Empirical frequencies of the first position of the output vector

Solution to Problem 1.17: First take $n = 0$, so that $I := I_0 = \int_0^\infty \left(x^2 + a\right)^{-1} dx$. Trial and error shows that $u = a/\left(x^2 + a\right)$ is a good substitution, with

$$x = a^{1/2}\left(\frac{1-u}{u}\right)^{1/2}, \quad dx = -a^{1/2}\frac{1}{2}\left(\frac{1-u}{u}\right)^{-1/2}\frac{1}{u^2}\,du,$$

so that

$$\begin{aligned} I &= -\int_1^0 \frac{u}{a}a^{1/2}\frac{1}{2}\left(\frac{1-u}{u}\right)^{-1/2}\frac{1}{u^2}du = \frac{a^{-1/2}}{2}\int_0^1 u^{-1/2}(1-u)^{-1/2}\,du \\ &= \frac{a^{-1/2}}{2}B\left(\frac{1}{2},\frac{1}{2}\right) = \frac{a^{-1/2}}{2}\frac{\Gamma(1/2)\,\Gamma(1/2)}{\Gamma(1)} = \frac{\pi}{2\sqrt{a}}. \end{aligned}$$

Exchanging derivative and integral,

$$\frac{dI}{da} = \int_0^\infty \frac{d}{da}\frac{1}{x^2+a}dx = -\int_0^\infty \left(x^2+a\right)^{-2} dx,$$

and, more generally,

$$\frac{d^nI}{da^n} = (-1)^n\, n! \int_0^\infty \left(x^2+a\right)^{-(n+1)} dx.$$

Now using $I = (\pi/2)\,a^{-1/2}$, it is easy to see that

$$\frac{d^nI}{da^n} = \frac{\pi}{2}(-1)^n\left(\frac{1}{2}\frac{3}{2}\frac{5}{2}\cdots\frac{2n-1}{2}\right)a^{-(2n+1)/2}$$

```
function y=permvec_pure(N)
y = permvec(N); cc=(1:N)';
olap = (cc == y);   % compute a boolean vector where overlap occurred
if sum(olap)>0
  oldy = y;         % just keep this in case we want to print the results
  loc=find(olap);   % loc is vector of locations where overlap occurred
  for i=1:sum(olap) % switch the offending entry with another element
    r=loc(i);
    while r==loc(i) % get another element
      r=unidrnd(N);
    end
    temp = y(r); y(r) = y(loc(i)); y(loc(i)) = temp;
  end
  if 1==1 % show the changes that had to be made
    loc
    check = [(1:N)' oldy y]
  end
end
```

Program Listing S-1.2: Output vector y is a permutation of vector $(1, 2, \ldots, N)$ with no 'coincidences'

so that, equating the two expressions,

$$\int_0^\infty \frac{1}{(x^2+a)^{n+1}}\,dx = \frac{\pi}{2}\left(\frac{1}{2}\frac{3}{4}\frac{5}{6}\cdots\frac{2n-1}{2n}\right)a^{-(2n+1)/2}.$$

Solution to Problem 1.18: From the hint, let

$$x = \frac{At+B}{Ct+D}, \quad \text{so that} \quad t = \frac{B-Dx}{Cx-A}.$$

When $x=0$, $t=0$ if $A \neq 0$ and $B=0$. When $x=1$, $t=1$ if $-(C-A) = D-B$, or $A = C+D$. That is, let

$$x = \frac{(C+D)\,t}{Ct+D}, \quad \text{with} \quad dx = \frac{(C+D)\,D}{(Ct+D)^2}dt.$$

Thus, with $1-x = D(1-t)\ /\ (Ct+D)$,

$$\begin{aligned} I &= \int_0^1 \frac{\left(\frac{(C+D)t}{Ct+D}\right)^{a-1}\left(\frac{D(1-t)}{Ct+D}\right)^{b-1}}{\left(\frac{(C+D)t}{Ct+D}+k\right)^{a+b}} \frac{(C+D)\,D}{(Ct+D)^2}\ dt \\ &= \int_0^1 \frac{t^{a-1}(1-t)^{b-1}(C+D)^a D^b}{((C+D)\,t+k\,(Ct+D))^{a+b}}dt. \end{aligned}$$

Some trial and error reveals that the constraint $C+D = -Ck$ leads to the simplification of the denominator term as

$$(C+D)\,t + k\,(Ct+D) = -Ckt + kCt + kD = kD = -Ck\,(k+1)\,,$$

so that, substituting and simplifying,

$$I = \int_0^1 t^{a-1}(1-t)^{b-1}\frac{1}{k^b(1+k)^a}dt,$$

which gives the answer.

Solution to Problem 1.19: Using integration by parts,

$$\begin{aligned}\int_{-1}^{1}(x^2-1)^j\,dx &= \int_{-1}^{1}(x-1)^j(x+1)^j\,dx\\ &= \int_{-1}^{1}(x-1)^j\,d\left(\frac{(x+1)^{j+1}}{j+1}\right)\\ &= \left[\frac{1}{j+1}(x-1)^j(x+1)^{j+1}\right]_{-1}^{1} - \int_{-1}^{1}\frac{j}{j+1}(x-1)^{j-1}(x+1)^{j+1}\,dx\\ &=(-1)\frac{j}{j+1}\int_{-1}^{1}(x-1)^{j-1}(x+1)^{j+1}\,dx.\end{aligned}$$

Repeating this,

$$\begin{aligned}&=(-1)\frac{j}{j+1}\int_{-1}^{1}(x-1)^{j-1}\,d\left(\frac{(x+1)^{j+2}}{j+2}\right)\\ &=-\left[\frac{j}{(j+1)(j+2)}(x-1)^{j-1}(x+1)^{j+2}\right]_{-1}^{1}\\ &\quad+(-1)^2\int_{-1}^{1}\frac{j(j-1)}{(j+1)(j+2)}(x-1)^{j-2}(x+1)^{j+2}\,dx\\ &=(-1)^2\frac{j(j-1)}{(j+1)(j+2)}\int_{-1}^{1}(x-1)^{j-2}(x+1)^{j+2}\,dx\\ &=\ \vdots\\ &=(-1)^j\frac{j!}{(2j)!/j!}\int_{-1}^{1}(x+1)^{2j}\,dx\\ &=(-1)^j\frac{1}{\binom{2j}{j}}\left[\frac{(x+1)^{2j+1}}{2j+1}\right]_{-1}^{1} = \frac{(-1)^j2^{2j+1}}{\binom{2j}{j}(2j+1)}.\end{aligned}$$

Solutions to Chapter 2: Probability Spaces and Counting

Solution to Problem 2.1:

a) This follows directly from property (iv).

b) As $B = A \cup \bar{A}B$, De Morgan's law gives $\bar{B} = \bar{A} \cap \left(A \cup \bar{B}\right) = \bar{A}\bar{B}$ so that $\bar{B} \subset \bar{A}$.

c) If $A \cap B$ implies C, then $A \cap B \subset C$. From De Morgan's law and the previous question, this is the same as $\bar{C} \subset \bar{A} \cup \bar{B}$. The result now follows from (ii) and (vi), i.e.,

$$\Pr\left(\bar{C}\right) \leq \Pr\left(\bar{A} \cup \bar{B}\right) \leq \Pr\left(\bar{A}\right) + \Pr\left(\bar{B}\right).$$

Solution to Problem 2.2:

a) $\Pr(A \cup B) \geq \Pr(A)$ and $\Pr(A \cup B) \geq \Pr(B)$ or

$$\Pr(A \cup B) \geq \max\{\Pr(A), \Pr(B)\} = 3/4.$$

b) Similarly, the upper bound is $\Pr(AB) \leq \min\{\Pr(A), \Pr(B)\} = 3/8$. From Bonferroni's equality $\Pr(AB) \geq \Pr(A) + \Pr(B) - 1 = 1/8$.

Solution to Problem 2.3:

(i) $\Pr(\emptyset) = 0$. Define $A_i = \emptyset \; \forall i$ so that $\emptyset = \bigcup_{i=1}^{\infty} A_i$ and, from additivity,

$$\Pr(\emptyset) = \Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i) = \sum_{i=1}^{\infty} \Pr(\emptyset),$$

which is only satisfied for $\Pr(\emptyset) = 0$.

(ii) If $A \subset B$, then $\Pr(A) \leq \Pr(B)$. Decompose event B as $B = A \cup \bar{A}B$ and note that A and $\bar{A}B$ are disjoint. From the finite additivity property, $\Pr(B) = \Pr(A) + \Pr\left(\bar{A}B\right)$ and, from nonnegativity of $\Pr(\cdot)$, $\Pr\left(\bar{A}B\right) \geq 0$, so that $\Pr(A) \leq \Pr(B)$.

(iii) $\Pr(A) \leq 1$. From (ii) with $A \subset \Omega$, $\Pr(A) \leq \Pr(\Omega) = 1$.

(iv) $\Pr(A^c) = 1 - \Pr(A)$. As A^c and A are disjoint but exhaust Ω, from $\Pr(\Omega) = 1$ and finite additivity,

$$1 = \Pr(\Omega) = \Pr(A \cup A^c) = \Pr(A) + \Pr(A^c).$$

(v) $\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \Pr(A_i)$. Define disjoint events

$$B_1 = A_1 \quad \text{and} \quad B_i = A_i \setminus (A_1 \cup A_2 \cup \cdots \cup A_{i-1}), \quad i > 1,$$

so that, from additivity, $\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \Pr\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \Pr(B_i)$. From (ii), $\Pr(B_i) \leq \Pr(A_i)$, $i > 1$, so that $\sum_{i=1}^{\infty} \Pr(B_i) \leq \sum_{i=1}^{\infty} \Pr(A_i)$, yielding the result.

(vi) $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$. By decomposing $A \cup B$ into three disjoint sets,

$$A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$$

or, from countable additivity,

$$\Pr(A \cup B) = \Pr(A \setminus B) + \Pr(B \setminus A) + \Pr(A \cap B) . \tag{S-2.1}$$

Similarly for A and B, $A = (A \cap B) \cup (A \setminus B)$ and $B = (A \cap B) \cup (B \setminus A)$, so that countable additivity and (S-2.1) imply that

$$\begin{aligned} \Pr(A) + \Pr(B) &= \Pr(A \cap B) + \Pr(A \setminus B) + \Pr(A \cap B) + \Pr(B \setminus A) \\ &= \Pr(A \cup B) + \Pr(A \cap B) . \end{aligned}$$

Subtracting $\Pr(A \cap B)$ from both sides yields the result.

Solution to Problem 2.4: Use of De Morgan's law and Poincaré's theorem gives

$$\begin{aligned} \Pr\left(\bigcap_{i=1}^{n} A_i\right) &= 1 - \Pr\left(\bigcup_{i=1}^{n} A_i^c\right) \\ &= 1 - \sum_{i=1}^{n} (-1)^{i+1} S_i \ , \quad S_j = \sum_{i_1 < \cdots < i_j} \Pr\left(A_{i_1}^c \cdots A_{i_j}^c\right) . \end{aligned}$$

But De Morgan's law implies

$$\Pr\left(A_{i_1}^c \cdots A_{i_j}^c\right) = 1 - \Pr\left(A_{i_1} \cup A_{i_2} \cdots \cup A_{i_j}\right) ,$$

so that, using the definition of R_j given in the problem,

$$S_j = \binom{n}{j} - R_j \ .$$

Substituting,

$$\Pr\left(\bigcap_{i=1}^{n} A_i\right) = 1 - \sum_{i=1}^{n} (-1)^{i+1} \left[\binom{n}{i} - R_i\right] = 1 - \sum_{i=1}^{n} (-1)^{i+1} \binom{n}{i} + \sum_{i=1}^{n} (-1)^{i+1} R_i \ ,$$

and the result follows, because of the binomial theorem (1.18), i.e.,

$$1 - \sum_{i=1}^{n} (-1)^{i+1} \binom{n}{i} = \sum_{i=0}^{n} (-1)^{i} \binom{n}{i} = \sum_{i=0}^{n} (-1)^{i} 1^{n-i} \binom{n}{i} = (-1+1)^n = 0.$$

Solution to Problem 2.5: The second equality in (2.24) follows from the countable additivity property, the fourth equality follows because the B_i are disjoint, the fifth equality is (A.4) and the sixth equality follows because the A_i are monotone increasing.

Solution to Problem 2.6:

a) Writing out the terms for when he discards the keys which do not work, we get

$$\frac{n-1}{n} \cdot \frac{n-2}{n-1} \cdot \cdots \cdot \frac{n-(k-1)}{n-(k-2)} \cdot \frac{1}{n-(k-1)} = 1/n.$$

b) If he does not discard, then the probability is $(n-1)^{k-1} n^{-k}$.

Solution to Problem 2.7:

a) The first person can sit in any of three cars, the same for the second, etc., so that $N = 3^6 = 729$.

b) Of the four remaining people, they have one of two cars to choose from, or $2^4 = 16$ possibilities, times $\binom{6}{2} = 15$ ways of selecting the two people who sit in the first car. Thus, the probability is $16 \times 15/729$.

Solution to Problem 2.8:

a) There are 3! ways of arranging the men in the first 3 chairs and 3! ways of arranging the women in chairs 4 through 6. Because we can switch the "block" of men and women, we have $2 \cdot (3!)^2$ ways, divided by the total number of ways of arranging 6 people in a row, 6!, yielding 0.1.

b) If we view the 3 men as a block, then we have 4 "objects" to arrange (the block of 3 men and 3 individual women), which can be done in 4! ways. Within the block of men, there are 3! arrangements, so the probability is $4! \cdot 3! \, / \, 6! = 0.2$.

c) Let both i_1, i_2, i_3 and j_1, j_2, j_3 be all permutations of $\{1, 2, 3\}$, let M_i denote the i^{th} man and W_j denote the j^{th} woman. Then there are 10 possible orderings, seen as follows. The most obvious is obtained by simply alternating the men and women, i.e.,

$$M_{i_1}, W_{j_1}, M_{i_2}, W_{j_2}, M_{i_3}, W_{j_3}$$

and, reversing it,

$$W_{i_1}, M_{j_1}, W_{i_2}, M_{j_2}, W_{i_3}, M_{j_3}.$$

In addition, (dropping the subscripts),

MWWMWM, MWMWWM, MMWWMW, WMWWMM,

WMMWMW, MWWMMW, WMWMMW and WMMWWM.

For each, there are $3! \cdot 3!$ different ways of arranging the men and women, yielding a probability of $10 \cdot (3!)^2 \, / \, 6! = 0.5$. Have we missed any combinations?

Solution to Problem 2.9: Take $b = 4$ for simplicity. Imagine placing the 4 beans into the 4 pieces of cake as follows. The first bean can go anywhere. The second has a 3/4 chance of getting into a piece of cake without a bean. The next bean has a 2/4 chance, the last a 1/4 chance, giving

$$\frac{3}{4}\frac{2}{4}\frac{1}{4} = \frac{4}{4}\frac{3}{4}\frac{2}{4}\frac{1}{4} = \frac{b!}{b^b}.$$

The suggestion $K_{b,b}^{(1)} / K_{b,b}^{(0)}$ is faulty because it assumes that all of the $K_{b,b}^{(0)}$ arrangements are equally likely, which they are not.

Solution to Problem 2.10: Let B (and G) be the events that you have at least one boy (girl). Then, from De Morgan's law,

$$\Pr(B \cup G) = 1 - \Pr(\overline{B \cup G}) = 1 - \Pr(\overline{B} \cap \overline{G}) = 1 - \Pr(\overline{B}) - \Pr(\overline{G}),$$

where the last equality follows because $\overline{B}$ and $\overline{G}$ are mutually exclusive, assuming $n > 0$. Assuming the probability of either gender is 0.5, this is

$$1 - \left(\frac{1}{2}\right)^n - \left(\frac{1}{2}\right)^n = 1 - \frac{1}{2^{n-1}}.$$

Solving, $n \geq 1 - \ln(1-\alpha) / \ln 2$.

For example, with $\alpha = 0.5$, $n \geq 2$, while for $\alpha = 0.95$, $n \geq 6$ and, for $\alpha = 0.999$, $n \geq 11$.

Solution to Problem 2.11: Directly,

$$\Pr\left(1^{\text{st}} \text{ try}\right) + \cdots + \Pr\left(5^{\text{th}} \text{ try}\right) = \frac{1}{8} + \frac{70}{80}\frac{10}{79} + \frac{70}{80}\frac{69}{79}\frac{10}{78} + \frac{70}{80}\frac{69}{79}\frac{68}{78}\frac{10}{77} + \frac{70}{80}\frac{69}{79}\frac{68}{78}\frac{67}{77}\frac{10}{76}$$

or $\binom{80}{5}^{-1} \sum_{i=1}^{5} \binom{10}{i}\binom{70}{5-i}$ or $1 - \Pr(\text{no underweight found}) = 1 - \binom{10}{0}\binom{70}{5}/\binom{80}{5} = 0.49655$.

Solution to Problem 2.12: There are $\binom{3}{1}\binom{97}{3}$ ways of selecting 1 of the three winning tickets and 3 of the 97 non–winning tickets; dividing this by the total number of ways of choosing 4 tickets out of the set of 100, or $\binom{100}{4}$, gives the probability of exactly one winning ticket. As having 1, 2, or 3 winning tickets are mutually exclusive events, probabilities add. Thus

$$p = \frac{\binom{3}{1}\binom{97}{3} + \binom{3}{2}\binom{97}{2} + \binom{3}{3}\binom{97}{1}}{\binom{100}{4}} = \frac{941}{8085} \approx 0.1164.$$

In most problems involving "at least", it is easier to work with the complement, which in this case is that no winning tickets are drawn. The probability of this is easily seen to be

$$q = 1 - p = \frac{\binom{3}{0}\binom{97}{4}}{\binom{100}{4}}.$$

An alternative method of solution is as follows. Instead of imagining choosing 4 tickets from the set of 100, imagine that the winning tickets are not yet determined, and are first separated into a group of four, which are yours, and the remaining 96. Now, the characteristic of "winning" is applied randomly to three of the 100 tickets. There are $\binom{4}{1}\binom{96}{2}$ ways of applying it to one of the four purchased tickets and two of the other 96, etc., so that

$$p = \frac{\binom{4}{1}\binom{96}{2} + \binom{4}{2}\binom{96}{1} + \binom{4}{3}\binom{96}{0}}{\binom{100}{3}} = 1 - \frac{\binom{96}{3}\binom{4}{0}}{\binom{100}{3}},$$

where the calculation involving the complement is also shown.

Now consider using Poincaré's theorem (2.11). First let

$$A_i = \{\text{you purchase winning ticket } i\}, \quad i = 1, 2, 3.$$

Then

$$\begin{aligned} p &= \Pr(A_1 \cup A_2 \cup A_3) \\ &= \binom{3}{1}\Pr(A_1) - \binom{3}{2}\Pr(A_1A_2) + \Pr(A_1A_2A_3) \\ &= 3 \cdot \frac{4}{100} - 3 \cdot \frac{4}{100}\frac{3}{99} + \frac{4}{100}\frac{3}{99}\frac{2}{98} = \frac{941}{8085} \approx 0.1164. \end{aligned}$$

Alternatively, let $B_i = \{\text{ticket } i \text{ is a winning ticket}\}$, $i = 1, 2, 3, 4$. Then

$$\begin{aligned} p &= \Pr(B_1 \cup B_2 \cup B_3 \cup B_4) \\ &= \binom{4}{1}\Pr(B_1) - \binom{4}{2}\Pr(B_1B_2) + \binom{4}{3}\Pr(B_1B_2B_3) - \Pr(B_1B_2B_3B_4) \\ &= 4 \cdot \frac{3}{100} - 6 \cdot \frac{3}{100}\frac{2}{99} + 4 \cdot \frac{3}{100}\frac{2}{99}\frac{1}{98} + 0. \end{aligned}$$

Solution to Problem 2.13: If $n < r$, then $p = 1$. For $n \geq r$, letting event A_i be that urn i is empty, $i = 1, \ldots, r$, it follows from Poincaré's theorem that

$$\begin{aligned} p &= 1 - \Pr\left(\bigcup_{i=1}^{r} A_i\right) \\ &= 1 - \sum_{i=1}^{r}\Pr(A_i) + \sum_{i=1}^{r}\sum_{j=i+1}^{r}\Pr(A_iA_j) - \cdots + (-1)^k \Pr(A_1A_2\cdots A_r), \end{aligned} \tag{S-2.2}$$

where $\Pr(A_i) = (1-p_i)^n$, $\Pr(A_iA_j) = (1-p_i-p_j)^n$, $i \neq j$, etc. Because the p_i are distinct, it appears as though little simplification is possible. There is, however, a recursive expression available which is computationally less burdensome than (S-2.2); see, e.g., Ross (1997, pp. 123-124).

Solution to Problem 2.14:

a) Possible outcomes are gb, gb, gg and bb with respective probabilities $p(1-p)$, $(1-p)p$, p^2 and $(1-p)^2$. Thus, the probability of at least one girl is $p^2 + 2p(1-p) = p(2-p)$.

b) The probability of having exactly k girls with order relevant is $p^k(1-p)^{n-k}$; it is easy to see that there are $n-k+1$ ways in which the k girls are all in a row. Thus, with $i = j - k$,

$$\begin{aligned} P(p,n,k) &= \sum_{j=k}^{n} (n-j+1)\, p^j (1-p)^{n-j} \\ &= p^k (1-p)^{n-k} \sum_{i=0}^{n-k} (n-i-k+1) \left(\frac{p}{1-p}\right)^i, \end{aligned}$$

which can also be expressed as

$$P(p,n,k) = p^k (1-p)^{n-k} \sum_{i=0}^{n-k} \sum_{j=0}^{i} \left(\frac{p}{1-p}\right)^j.$$

Then, with $a = p/(1-p)$,

$$\begin{aligned} \sum_{i=0}^{n-k} \sum_{j=0}^{i} a^j &= \sum_{i=0}^{n-k} \frac{1-a^{i+1}}{1-a} = \frac{1}{1-a}\left((n-k+1) - \sum_{i=0}^{n-k} a^{i+1}\right) \\ &= \frac{1}{1-a}\left((n-k+1) - \frac{a-a^{n-k+2}}{1-a}\right) \\ &= \frac{1-p}{1-2p}\left((n-k+1) - \left(\frac{p}{1-2p}\right)\left(1-\left(\frac{p}{1-p}\right)^{n-k+1}\right)\right) \end{aligned}$$

so that, for $p \neq 1/2$,

$$P(p,n,k) = p^k \frac{(1-p)^{n-k+1}}{1-2p}\left(n-k+1-\frac{p}{1-2p}\left(1-\left(\frac{p}{1-p}\right)^{n-k+1}\right)\right).$$

Similarly, for $p = 1/2$,

$$\begin{aligned} P\left(\frac{1}{2}, n, k\right) &= 2^{-n} \sum_{j=k}^{n} (n-j+1) \\ &= 2^{-n}\left[(n+1)(n-k+1) - \frac{1}{2}\left(n(n+1) - k(k-1)\right)\right] \\ &= 2^{-(n+1)} (n-k+2)(n-k+1). \end{aligned}$$

For $n = 2$ and $k = 1$, Maple admirably shows that $P(p,2,1) = p(2-p)$. For $n = 7$ and $k = 3$, Maple gives

$$P(p,7,3) = p^3\left(3p^4 - 12p^3 + 21p^2 - 16p + 5\right).$$

Note that, in both cases, the special formula for $p = 1/2$ is not needed. Also, $P(1/2, 7, 3) = 15/128 \approx 0.117$.

The left side of Figure S-2.1 plots $P(p,7,3)$ versus p (dashed line with $\times$ markings) along with the result from Example 2.7 (solid line with $\circ$ markings), i.e., the probability of getting at least 3 girls in a row when having 7 children (but ALL the girls do NOT have to be in a row, just 3 of them).

Solution to Problem 2.15:

a) As there is only one correct permutation, the probability is $1/n!$.

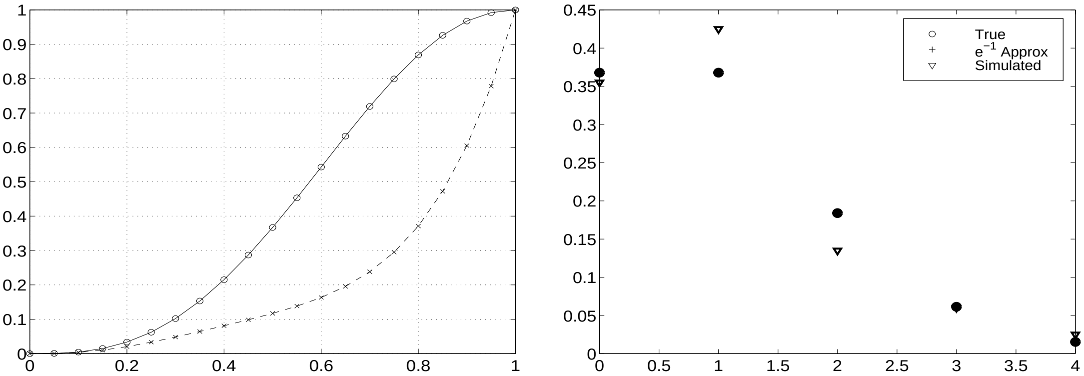


Figure S-2.1: Left: Probability $P(p, 7, 3)$ (dashed) versus p. Right: Output from calling `permvecsim(20,200)`. Because the approximation using e^{-1} is so good for $N = 20$, the circles and plus signs cannot be distinguished in the printed version.

b) This is just (2.18).

c) $1 - \Pr(k = 0) \approx 1 - e^{-1} = 0.632$ for large n.

Solution to Problem 2.16: Program S-2.1 does this. Sample graphical output is given in the right side of Figure S-2.1, obtained from calling `permvecsim(20,200)`.

Solution to Problem 2.17: Define C to be the event such that the hand contains the Ace and King of Clubs (♣). Likewise for D (Diamonds, ♢), H (Hearts, ♡) and S (Spades, ♠). We need to compute $\Pr(C \cup D \cup H \cup S)$ for which we can use (2.11). For the 4 events, we write

$$\begin{aligned}\Pr(C \cup D \cup H \cup S) = {} & \Pr(C) + \Pr(D) + \Pr(H) + \Pr(S) \\ & - \Pr(CD) - \Pr(CH) - \Pr(CS) \\ & - \Pr(DH) - \Pr(DS) - \Pr(HS) \\ & + \Pr(CDH) + \Pr(CDS) + \Pr(CHS) \\ & + \Pr(DHS) - \Pr(CDHS),\end{aligned}$$

and the event C can occur in any of $\binom{50}{11}$ ways, because we know 2 of the cards and must pick the remaining 11 from the remaining 50. Event CD can occur in $\binom{48}{9}$ ways, because we know 4 cards and must pick the remaining 9 from the remaining 48. Likewise for event CDH, which can occur in any of $\binom{46}{7}$ ways, and event $CDHS$, which can occur in any of $\binom{44}{5}$ ways. Notice that event D can occur in as many ways as event C and similarly for the rest, so that the probability is given by

$$\frac{4\binom{50}{11} - 6\binom{48}{9} + 4\binom{46}{7} - \binom{44}{5}}{\binom{52}{13}} = \binom{52}{13}^{-1} \sum_{i=1}^{4} (-1)^{i+1} \binom{4}{i} \binom{52-2i}{13-2i} = \frac{9895443}{45023650} = 0.21978.$$

Solution to Problem 2.18: The second child's sex is (presumably) independent of that of the first child, so it is just Pr(boy) = 0.5.

Solution to Problem 2.19: If B stands for boy and G for girl, then the relevant or conditional sample space is $\{B, B\}, \{G, B\}$ or $\{B, G\}$, so that the probability of 2 boys is 1/3.

Solution to Problem 2.20: Let the event R_i denote the event that the first red ball is appears on the i^{th} draw. *If you draw first*, then your chances of winning are

$$\sum_{i=1,3,5,7} \Pr(R_i),$$

```
function [simprob, trueprob, approxprob] = permvecsim(N,sim)
record=zeros(1,sim); cc=(1:N)';
for i=1:sim
  y = permvec(N);
  record(i)=sum(cc == y);
end
maxr = max(record); g=hist(record,maxr+1); simprob=g/sim;
trueprob=zeros(1,maxr+1); approxprob=trueprob;
for m=0:maxr
  v=2:(N-m); v=fact(v); v=1./v;
  pm=(-1).^(2:(N-m)); fm = fact(m);
  trueprob(m+1)=sum(v.*pm)/fm;
  approxprob(m+1) = exp(-1)/fm;
end

plot(0:maxr,trueprob,'linewidth',4,'linestyle','o', ...
     'color','r','marker','o'), hold on
plot(0:maxr,approxprob,'linewidth',3,'linestyle','o', ...
     'color','g','marker','+')
plot(0:maxr,simprob,'linewidth',2,'linestyle','o', ...
     'color','b','marker','v'), hold off
legend('True','e^{-1} Approx','Simulated'), set(gca,'fontsize',16)

function f=fact(x), f=round( gamma(x+1) );
```

Program Listing S-2.1: Simulates the number of coincidences and compares the empirical frequencies to the true probabilities (2.18) and to the approximation $e^{-1}/m!$. Assumes program `permvec` exists, which is given in Listing 1.2

because the events R_i, $i = 1, \ldots, 8$ are disjoint. The probability that your friend wins is $\sum_{i=2,4,6,8} \Pr(R_i)$, but, as the events R_i, $i = 1, \ldots, 8$ also partition the entire sample space, we know that

$$\sum_{i=2,4,6,8} \Pr(R_i) = 1 - \sum_{i=1,3,5,7} \Pr(R_i).$$

We have

$$\Pr(R_1) = \frac{3}{10}, \quad \Pr(R_3) = \frac{7}{10}\frac{6}{9}\frac{3}{8}, \quad \Pr(R_5) = \frac{7}{10}\frac{6}{9}\frac{5}{8}\frac{4}{7}\frac{3}{6}, \quad \Pr(R_7) = \frac{7}{10}\frac{6}{9}\frac{5}{8}\frac{4}{7}\frac{3}{6}\frac{2}{5}\frac{3}{4},$$

so that

$$\sum_{i=1,3,5,7} \Pr(R_i) = \frac{7}{12} > \frac{1}{2},$$

so it is advantageous to draw first.

Solution to Problem 2.21: If you see a red dot, then the relevant, or conditional sample space is 3 events: either you see *one* side of the red-red card, the *other* side of the red-red card, or the red dotted side of the red-black card, so that the probability that the other side has a black dot is $\frac{1}{3}$. More formally, the sample space could be written as follows. As there are 6 sides which could be displayed, there are 6 events that can happen. With R for red, and B for black, we have $\{RR\}, \{RR\}, \{RB\}, \{BR\}, \{BB\}, \{BB\}$, where the first letter corresponds to the visible side of the card and the second letter to the covered side. If you see a red dot, that means the events starting with B, namely $\{BB\}$, $\{BB\}$ and $\{BR\}$, are not possible. The reduced sample space is the 3 events $\{RR\}, \{RR\}$ and $\{RB\}$, so the probability of the covered side being black is clearly $1/3$.

Solution to Problem 2.22:

a) Because games are independent,

$$\Pr(A) = p_1 + q_1 q_2 p_1 + q_1 q_2 q_1 q_2 p_1 + \cdots = p_1 \sum_{i=0}^{\infty} (q_1 q_2)^i = \frac{p_1}{1 - q_1 q_2}$$

and, likewise, $\Pr(B) = p_2 q_1 \sum_{i=1}^{\infty} (q_1 q_2)^{i-1} = \frac{p_2 q_1}{1 - q_1 q_2} = 1 - \Pr(A)$.

b) This just means that $p_1 = p_2 q_1$ or $p_2 = p_1 / (1 - p_1)$. Condition $0 < p_2 < 1$ implies $p_1 < 1 - p_1$ or that $p_1 < 0.5$. As $p_1 \to 0.5$, for a fair game, p_2 must approach one, which adjusts for the fact that person A plays first.

Solution to Problem 2.23: Both sides are equal to $\binom{m+b-1}{b-1}$. To see this for the lhs, in (2.3), replace r by b and n by m. For the rhs, in (1.32), replace n by k and r by m.

Solution to Problem 2.24: In this case, S_j reduces to

$$S_j = \sum_{i_1 < \cdots < i_j} \Pr\left(A_{i_1} \cdots A_{i_j}\right) = \sum_{i_1 < \cdots < i_j} p^j = \binom{n}{j} p^j,$$

so that

$$p_{m,n} = \sum_{i=m}^{n} (-1)^{i-m} \binom{i}{i-m} \binom{n}{i} p^i.$$

This gives $p_{2,4} = 6p^2 - 3 \cdot 4p^3 + 6p^4$, while in terms of complements, $p_{2,4} = 6p^2 (1-p)^2$, which are clearly equal. Similarly, $P_{2,4} = 6p^2 - 2 \cdot 4p^3 + 3p^4$ or, in terms of complements, $P_{2,4} = 1 - (1-p)^4 - 4p(1-p)^3$, which are easily shown to be equal.

Solution to Problem 2.25: We have:

1. For 3 cells empty: There are $\binom{R}{1} = 4$ possible ways of picking the nonempty cell.
2. For 2 cells empty: If the balls are indistinguishable, there are $\binom{R}{2} = 6$ ways to pick the two empty cells. For the nonempty cells, there are $K_{n,r-2}^{(1)} = \binom{n-1}{r-3} = 2$ distributions, either a cell has one ball, the other two; or visa versa. Factoring in ball distinguishability, there is $\binom{3}{1}$ ways of choosing "the lone ball".
3. For 1 cell empty: If the balls are indistinguishable, there are $\binom{R}{1} = 4$ ways to pick the empty cell. For the nonempty cells, there is clearly only one possible arrangement, i.e., each cell has one ball (Or, $K_{n,r-1}^{(1)} = \binom{n-1}{r-2} = 1$.) Factoring in ball distinguishability, this is multiplied by $n! = 6$.
4. No cells empty cannot occur.

This gives $4 + 6 \cdot 2 \cdot 3 + 4 \cdot 6 = 64$ possibilities. If both balls and cells are indistinguishable, then one of three configurations can arise, $\{\bullet \bullet \bullet \mid \cdot \mid \cdot \mid \cdot\}$, or $\{\bullet\bullet \mid \bullet \mid \cdot \mid \cdot\}$ or $\{\bullet \mid \bullet \mid \bullet \mid \cdot\}$, with respective probabilities 4/64, 36/64 and 24/64.

Solutions to Chapter 3: Symmetric Spaces and Conditioning

Solution to Problem 3.1:

a) Assume $t = 5$ and that, contrary to the initial setup, the order of the throws is relevant. By the nature of the experiment, all 6^5 outcomes are equally likely. Say sides 1 and 2 are deemed not to occur and denote throw i as w_i. Then, all possible outcomes of the five throws such that a 1 or 2 never occurs could be listed in table:

w_1	w_2	w_3	w_4	w_5
3	3	3	3	3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Instead, we could make a table showing which of the w_i were 3, 4, etc., as

3	4	5	6
w_1	w_2	w_3	w_4, w_5

As all four sides must be covered, one throw will result in a side which already occurred, with 4 possibilities. If an extra column is added to the above table which picks up the "redundant side", then there will be one w_i per box, with 5! ways of permuting them. As we are ultimately not concerned with order, this gets divided by two to account for the redundancy involved for the side with two occurrences. Finally, from symmetry, the same argument applies for any 2 chosen sides which should not occur, i.e., for all $\binom{6}{2} = 15$ choices. Thus,

$$p_5 = \Pr(\text{exactly 2 sides do not occur}) = \frac{15 \cdot 5! \cdot 4 \,/\, 2}{6^5} = \frac{25}{54}.$$

For $t = 6$, the "redundant" two throws are either such that two sides each occur twice, or one side occurs three times. Thus,

$$p_6 = \frac{\binom{6}{2} \cdot 6! \cdot \binom{4}{2} \,/\, 2^2}{6^6} + \frac{\binom{6}{2} \cdot 6! \cdot \binom{4}{1} \,/\, 3!}{6^6} = \frac{325}{648}.$$

b) Let events $B_i = \{\text{side } i \text{ does not appear}\}$, $i = 1, 2, \ldots, 6$, so that $\Pr(B_i) = (5/6)^t$ and

$$\Pr(B_1 \cdots B_j) = \binom{6}{j} (1 - j/6)^t .$$

Let E_2 be the event that neither a 1 nor 2 occurs but 3,4,5,6 all occur at least once and let $L_i = B_1 B_2 \cdots B_i$. Note that, for example, L_2 is the event that neither 1 nor 2 occurs, irrespective of the other numbers. Then L_2 can be written as the disjoint union of E_2 and $L_2 \cap (B_3 \cup \cdots \cup B_6)$, i.e.,

$$\Pr(L_2) = \Pr(E_2) + \Pr((L_2 \cap B_3) \cup \cdots \cup (L_2 \cap B_6)),$$

so that, from Poincaré's theorem and the exchangeability of the B_i,

$$\begin{aligned}
\Pr(E_2) &= \Pr(L_2) - \binom{4}{1}\Pr(L_2 \cap B_3) + \binom{4}{2}\Pr(L_2 \cap B_3 \cap B_4) - \cdots \\
&= \sum_{i=2}^{6} (-1)^{i-2} \binom{4}{i-2} \Pr(L_i) \\
&= \sum_{i=2}^{6} (-1)^{i-2} \binom{4}{i-2} (1 - i/6)^t .
\end{aligned}$$

As there are $\binom{6}{2} = 15$ ways of picking two sides,

$$p_t = 15 \sum_{i=2}^{6} (-1)^{i-2} \binom{4}{i-2} (1 - i/6)^t = \sum_{i=2}^{6} (-1)^{i-2} \binom{i}{2}\binom{6}{i} (1 - i/6)^t ,$$

because

$$\frac{6!}{2!4!}\frac{4!}{(i-2)!\,(6-i)} = \frac{i!}{2!\,(i-2)!}\frac{6!}{i!\,(6-i)} = \binom{i}{2}\binom{6}{i}.$$

Solution to Problem 3.2:

a) From basic principles,

$$p_{6n}(n) = \frac{\binom{6n}{n,n,n,n,n,n}}{6^{6n}} = \frac{(6n)!}{(n!)^6\, 6^{6n}} \approx \frac{(2\pi)^{1/2}(6n)^{6n+1/2} e^{-6n}}{(2\pi)^{6/2} n^{6(n+1/2)} e^{-6n}\, 6^{6n}} = (2\pi)^{-5/2}\, 6^{1/2} n^{-5/2}.$$

The exact answer for $n = 3$ is

$$p_{18} = \frac{14889875}{1\,10199\,60576} = 1.35 \times 10^{-3}.$$

The relative percentage errors

$$100 \frac{(2\pi)^{-\frac{5}{2}}\sqrt{6} n^{-\frac{5}{2}} - \frac{(6n)!}{(n!)^6 6^{6n}}}{\frac{(6n)!}{(n!)^6 6^{6n}}}$$

for $n = 1, 2, \ldots, 20$ are computed to be

t	1	2	3	4	5	6	7	8	9	10
RPE	60	27	18	13	10.2	8.4	7.2	6.3	5.5	5.0
t	11	12	13	14	15	16	17	18	19	20
RPE	4.5	4.1	3.8	3.5	3.3	3.1	2.9	2.7	2.6	2.5

.

b) Clearly, if $t \le 17$, then $p_t = 0$. Otherwise, defining

$$B_i = \{\text{ side } i \text{ occurs zero or once or twice}\},$$

we have

$$p_t = \Pr(\text{all 6 sides occur at least twice}) = 1 - \Pr\left(\bigcup_{i=1}^{6} B_i\right).$$

Letting $N_i = \#$ of times side i occurs, event B_i can be decomposed into the three disjoint events $\{N_i = 0\}, \{N_i = 1\}$ and $\{N_i = 2\}$, so that

$$\Pr(B_i) = \left(\frac{5}{6}\right)^t + \binom{t}{1}\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^{t-1} + \binom{t}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^{t-2}.$$

Similarly, in abbreviated notation for the N_i, event B_1B_2 can be decomposed into the 3^2 disjoint events (grouped in 6 categories)

$$\begin{aligned}&\{00\}\,,\\&\{11\}\,,\\&\{22\}\,,\\&\{01\}\,,\{10\}\,,\\&\{02\}\,,\{20\}\,,\\&\{12\}\,,\{21\}\,,\end{aligned}$$

so that $\Pr(B_1B_2)$ is given by

$$\begin{aligned}=\;&\left(\frac{4}{6}\right)^t+\binom{t}{1,1,t-2}\left(\frac{1}{6}\right)^2\left(\frac{4}{6}\right)^{t-2}+\binom{t}{2,2,t-4}\left(\frac{1}{6}\right)^4\left(\frac{4}{6}\right)^{t-4}\\&+2\binom{t}{1}\left(\frac{1}{6}\right)\left(\frac{4}{6}\right)^{t-1}+2\binom{t}{2}\left(\frac{1}{6}\right)^2\left(\frac{4}{6}\right)^{t-2}+2\binom{t}{1,2,t-3}\left(\frac{1}{6}\right)^3\left(\frac{4}{6}\right)^{t-3}.\end{aligned}$$

Event $B_1B_2B_3$ can be decomposed into the 3^3 disjoint events (grouped in 10 categories)

1. $\{000\}\,,\quad \binom{3}{3,0,0}$ of them,
2. $\{111\}\,,\quad \binom{3}{0,3,0}$ of them,
3. $\{222\}\,,\quad \binom{3}{0,0,3}$ of them,
4. $\{001\}\,,\{010\}\,,\{100\}\,,\quad \binom{3}{2,1,0}$ of them,
5. $\{002\}\,,\{020\}\,,\{200\}\,,\quad \binom{3}{2,0,1}$ of them,
6. $\{011\}\,,\{101\}\,,\{110\}\,,\quad \binom{3}{1,2,0}$ of them,
7. $\{112\}\,,\{121\}\,,\{211\}\,,\quad \binom{3}{0,2,1}$ of them,
8. $\{122\}\,,\{212\}\,,\{221\}\,,\quad \binom{3}{0,1,2}$ of them,
9. $\{022\}\,,\{202\}\,,\{220\}\,,\quad \binom{3}{1,0,2}$ of them,
10. $\{012\}\,,\{102\}\,,\{120\}\,,\ \{021\}\,,\{201\}\,,\{210\}\,,\quad \binom{3}{1,1,1}$ of them,

so that $\Pr(B_1B_2B_3)$ is given by

$$\begin{aligned}=\;&\left(\frac{3}{6}\right)^t+\binom{t}{1,1,1,t-3}\left(\frac{1}{6}\right)^3\left(\frac{3}{6}\right)^{t-3}+\binom{t}{2,2,2,t-6}\left(\frac{1}{6}\right)^6\left(\frac{3}{6}\right)^{t-6}\\&+3\binom{t}{0,0,1,t-1}\left(\frac{1}{6}\right)^1\left(\frac{3}{6}\right)^{t-1}+3\binom{t}{0,0,2,t-2}\left(\frac{1}{6}\right)^2\left(\frac{3}{6}\right)^{t-2}\\&+3\binom{t}{0,1,1,t-2}\left(\frac{1}{6}\right)^2\left(\frac{3}{6}\right)^{t-2}+3\binom{t}{1,1,2,t-4}\left(\frac{1}{6}\right)^4\left(\frac{3}{6}\right)^{t-4}\\&+3\binom{t}{1,2,2,t-5}\left(\frac{1}{6}\right)^5\left(\frac{3}{6}\right)^{t-5}+3\binom{t}{0,2,2,t-4}\left(\frac{1}{6}\right)^4\left(\frac{3}{6}\right)^{t-4}\\&+6\binom{t}{0,1,2,t-3}\left(\frac{1}{6}\right)^3\left(\frac{3}{6}\right)^{t-3}.\end{aligned}\tag{S-3.1}$$

It is becoming clear that, for $\Pr(B_1 B_2 B_3)$, we need to consider all nonnegative integer solutions to $\sum_{i=1}^{3} x_i = 3$, for which there are $\binom{3+3-1}{3} = 10$, using (2.1). (Similarly, for $\Pr(B_1 B_2)$, there are six nonnegative integer solutions to $\sum_{i=1}^{3} x_i = 2$.) For each of the 10 solutions, there are $\binom{3}{x_1, x_2, x_3}$ combinations such that x_1 of the N_i are zero, x_2 of the N_i are one and x_3 of the N_i are two. Then, for each of the nonnegative integer solutions, a term in (S-3.1) can be determined as

$$\binom{3}{x_1, x_2, x_3}\binom{t}{S_1, S_2, S_3, t-S}\left(\frac{1}{6}\right)^{S}\left(\frac{3}{6}\right)^{t-S},$$

where $S = \sum_{i=1}^{3} S_i$, $(S_1, S_2, S_3) = \text{sort}(V)$ (ascending) and V is a vector consisting of x_1 zeros, x_2 ones and x_3 twos. In this way, the first element of each row of the 27 listed expressions above would be taken.

One problem with this formulation is that we would now need an algorithm which can list out the $\binom{3+3-1}{3}$ nonnegative integer solutions. This could be achieved by a double FOR loop:

```
FOR x1=0:3
  FOR x2=0:(3-x1)
    x3=3-x1-x2
    --- now use these x1 x2 and x3 ---
  END
END
```

This was just for the $\Pr(B_1 B_2 B_3)$ term, but the generalization to $\Pr(B_1 \cdots B_j)$ is now straightforward. In particular, we need the nonnegative integer solutions to $\sum_{i=1}^{3} x_i = j$, for which there are $\binom{j+3-1}{j} = \binom{j+2}{j} = (j+2)(j+1)/2$. For each of them, there are $\binom{j}{x_1, x_2, x_3}$ combinations such that x_1 of the N_i are zero, x_2 of the N_i are one and x_3 of the N_i are two. Then, for each of these nonnegative integer solutions, the terms in $\Pr(B_1 \cdots B_j)$ can be expressed as

$$\binom{j}{x_1, x_2, x_3}\frac{t!}{S_1! \cdots S_j!\,(t-S)!}\left(\frac{1}{6}\right)^{S}\left(\frac{6-j}{6}\right)^{t-S},$$

where $S = \sum_{i=1}^{j} S_i$, $(S_1, \ldots S_j) = \text{sort}(V)$ (ascending) and V is a vector consisting of x_1 zeros, x_2 ones and x_3 twos. The program to traverse the $(j+2)(j+1)/2$ solutions would look like

```
FOR x1=0:j
  FOR x2=0:(j-x1)
    x3=j-x1-x2
    --- now use these x1 x2 and x3 ---
  END
END
```

A computer program to compute p_t and also simulate the problem is given in Listing S-3.1. Figure S-3.1 plots p_t versus t with the inscribed circles indicating the result of simulating the result for each shown value of t using 5,000 replications. It was constructed with the following Matlab instructions:

```
pt=[];
for i=18:2:60
  pt = [pt sixes(i,0)];
end
plot(18:2:60,pt), hold on
sim=5000; sv=[]; for i=20:4:60; i, sv=[sv sixes(i,sim)]; end
plot(20:4:60,sv,'ro')
```

```
function pt = sixes(t,sim)
if sim>0
  count=0;
  for i=1:sim
    u=unidrnd(6,t,1);       % t throws of the die
    tab=tabulate(u);        % counts how many ones, how many twos, etc.
    f=tab(:,2)';            % just the part we need from the output
    count=count+all(f>=3);  % check if all sides were observed
  end                       %                 at least 3 times
  pt = count/sim;           % approximate the true probability
else % Compute exact probability using Poincare's theorem
  bigsum = 0;
  for j=1:5
    bigsum=bigsum + (-1)^(j+1) * c(6,j) * bigterm(j,t);
  end
  pt = 1-bigsum;
end

function bt = bigterm(j,t) % computes Pr(B_1 ... B_j)
bt=0;
for x1=0:j
  for x2=0:(j-x1)
    x3=j-x1-x2;
    f1 = fact(j) / fact(x1)  / fact(x2)  / fact(x3);
    thezeros=zeros(1,x1); theones=ones(1,x2); thetwos=2*ones(1,x3);
    V = [thezeros theones thetwos];
    Svec = sort(V); S=sum(Svec);
    denom=1; for i=1:j, denom=denom*fact(Svec(i)); end
    denom=denom*fact(t-S);
    f2 = fact(t) / denom;
    f3 = (1/6)^S * ((6-j)/6)^(t-S);
    bt = bt + f1 * f2 * f3;
  end
end

function f=fact(t) % simple, local, factorial function
  if t==0, f=1; else f=t*fact(t-1); end
```

Program Listing S-3.1: A fair, six–sided die is rolled t times. This computes the probability that all 6 sides occur at least 3 times. If `sim` > 0, then it simulates instead of computing the theoretical values

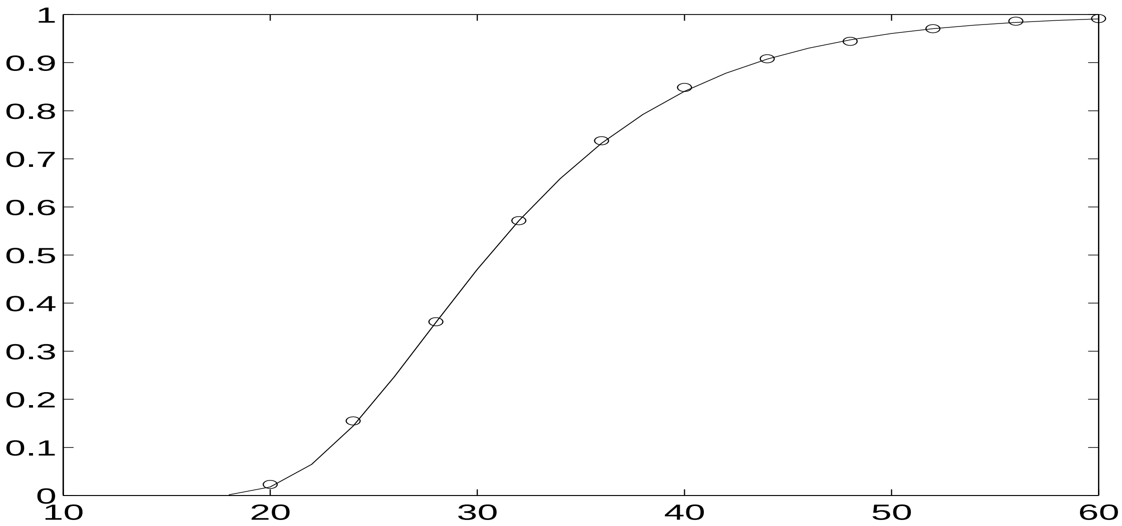


Figure S-3.1: Solid line is the exact probability that all 6 sides occur at least 3 times. Circles indicate result of simulation using 5,000 replications

Solution to Problem 3.3:

a) 11!, because, if they were in a row, there would be 12!, but the rotation of the table doesn't matter.

b) If each couple sits next to one another, we can imagine the 12 people as 6 objects to be placed around a table with 6 places, and there are 5! ways of doing this. As each couple can "switch chairs", the answer is $(2^6 \cdot 5!) \,/\, 11! = 1.924 \times 10^{-4}$.

Solution to Problem 3.4:

a) Assume A sits first at the empty table and then B takes a seat. Of the $n-1$ possible chairs, there are two which B could take such that A and B sit together, i.e., the probability is $2/(n-1$. Alternatively, and more generally applicable, treating A and B as a block the probability is $\frac{2(n-2)!}{(n-1)!} = \frac{2}{n-1}$.

b) Similar to the latter method in the previous question, i.e., treating [A,B] and [C,D] as blocks, $\frac{2^2(n-3)!}{(n-1)!} = \frac{4}{(n-2)(n-1)}$. This can be also obtained by using the law of total probability. Let AB be the event that A and B sit together, which is $2/(n-1)$. Then

$$\Pr(AB \cap CD) = \Pr(AB)\Pr(CD \mid AB).$$

Assume C is the third person to take a seat. To calculate $\Pr(CD \mid AB)$, we have to consider if C sits on either side of the pair AB (with probability $2/(n-2)$), or not. That is,

$$\Pr(CD \mid AB) = \frac{2}{n-2}\frac{1}{n-3} + \frac{n-4}{n-2}\frac{2}{n-3} = \frac{2}{n-2}$$

so that

$$\Pr(AB \cap CD) = \frac{2}{n-1}\frac{2}{n-2} = \frac{4}{(n-1)(n-2)}.$$

c) The two possibilities are not equally likely. Assume the first person to sit is a boy at the 12:00 position, as in Figure 3.9, followed by the second boy. In order to alternate, he has to sit at the 6:00 position whereas to not alternate, he can sit either at 3:00 or 9:00. The chances are double that they do not alternate, so the correct probability is 1/3. Another way of seeing this is to "open the table" into a row and consider the $4!/(2!2!) = 6$ possibilities, i.e., BBGG, BGBG, BGGB, GGBB, GBGB and GBBG. When "re-connecting", only 2 of the 6 combinations yield alternating sequences.

d) Because they must alternate, take each boy-girl pair as a block. These have $(n-1)!$ ways of being arranged. The first girl can be paired with any of n boys, the second with any of $n-1$, etc., yielding $(n-1)! \cdot n! \,/\, (2n-1)!$.

Solution to Problem 3.5:

a) Event $K = n$ means all couples are seated together. This is just (3.3) with $i = n$, i.e.,

$$\Pr(C_1 \cdots C_n) = 2^n \frac{(n-1)!}{(2n-1)!}.$$

For $K = n-1$, there are n ways of choosing the "lone" couple, i.e., the one couple not sitting together. Then, there are 2^{n-1} ways of shuffling the $n-1$ other couples internally and $n+1$ 'blocks' to arrange around the table. But this also counts the possibilities in which the lone couple is together, for which there are $2^{n-1}2(n-1)!$ such arrangements. Thus,

$$\Pr(K = n-1) = n\frac{2^{n-1}\left(n! - 2\,(n-1)!\right)}{(2n-1)!} = \frac{n!2^{n-1}\,(n-2)}{(2n-1)!},$$

which agrees with (3.4).

b) To compute the probability of no couples, the complement "at least one couple" can be used, i.e., from Poincaré's theorem and (3.3),

$$\begin{aligned}
\Pr(K=0) &= 1 - \Pr(C_1 \cup \ldots \cup C_n) \\
&= 1 - \sum_{i=1}^{n} (-1)^{i+1} \binom{n}{i} \frac{2^i (2n-i-1)!}{(2n-1)!} \\
&= \sum_{i=0}^{n} (-1)^{i} \binom{n}{i} \frac{2^i (2n-i-1)!}{(2n-1)!},
\end{aligned}$$

which agrees with (3.4).

For $k = 1$, there are n ways of choosing the lucky couple to sit together and this couple has 2 ways to sit. The other $2n-2$ people must be arranged in such a way that no other couple is seated together. The number of possibilities can be calculated in the same way as in the case $k = 0$, with the difference that now the $2n-2$ people will be arranged in a row, because their "ends" do not touch (they are separated by the lucky couple) and thus cannot be rotated. This yields, using $j = i+1$,

$$\begin{aligned}
\Pr(K=1) &= \frac{2n}{(2n-1)!} \left((2n-2)! - \sum_{i=1}^{n-1} (-1)^{i+1} \binom{n-1}{i} 2^i \,(2n-2-i)! \right) \\
&= 2n \left(\frac{1}{2n-1} - \sum_{j=2}^{n} (-1)^{j} \binom{n-1}{j-1} 2^{j-1} \frac{(2n-1-j)!}{(2n-1)!} \right) \\
&= n \left(\frac{2}{2n-1} + \sum_{j=2}^{n} (-1)^{j+1} \binom{n-1}{j-1} 2^{j} \frac{(2n-1-j)!}{(2n-1)!} \right) \\
&= n \sum_{j=1}^{n} (-1)^{j+1} \binom{n-1}{j-1} \frac{2^j \,(2n-1-j)!}{(2n-1)!}
\end{aligned}$$

which agrees with (3.4).

c) First observe that, if k is negative or $k > n$, then $C_{k,n} = 0$. For the starting values, if $n = 1$, then the couple must sit together, so that $C_{0,1} = 0$ and $C_{1,1} = 1$. If $n = 2$, then, denoting the first couple as A_1 and A_2 and the second as B_1 and B_2,

i) if $k = 0$, there are two possibilities, (starting at 12:00 and going clockwise),

$$A_1B_1A_2B_2, \quad \text{and} \quad A_1B_2A_2B_1\,;$$

ii) if $k = 2$, there are four possibilities,

$$A_1A_2B_1B_2,\ A_2A_1B_1B_2,\ A_1A_2B_2B_1,\ A_2A_1B_2B_1;$$

iii) the case $k = 1$ cannot occur.

Note how the rotation of the table reduces the number of possibilities. Summarizing,

$$C_{0,1} = 0,\ C_{1,1} = 1,\ C_{0,2} = 2, C_{1,2} = 0,\ C_{2,2} = 4.$$

Imagine the couples coming to the table one after another. For $n > 2$, consider how persons P_1 and P_2 making up the n^{th} couple were seated so that there are exactly k couples sitting together. Either:

i) There were exactly $k-1$ couples sitting together previous to the n^{th} couple's arrival and P_1 and P_2 sat down next to each other, but not "between any couples", i.e., such that they do not separate any other couple. Of the $2n-2$ spaces at the table, $k-1$ cannot be used, so they have

$$2\,(2\,(n-1)-(k-1)) = 2\,(2n-k-1)$$

ways to do this.

ii) There were exactly k couples sitting together previous to the n^{th} couple's arrival and either

1. persons P_1 and P_2 sat down together between a couple, with $2k$ ways, or
2. P_1 sat down not between any of the k couples in one of $2\,(n-1)-k$ ways and P_2 sat down not between any of the k couples and not next to P_1 in one of $2\,(n-1)+1-(k+2)$ ways. Multiplying gives $(2n-k-2)\,(2n-k-3)$ ways.

iii) There were exactly $k+1$ couples sitting together previous to the n^{th} couple's arrival; P_1 sits between one of the couples (with $k+1$ ways) and P_2 sits neither between any of the remaining k couples, nor next to P_1 (with $2\,(n-1)+1-(k+2)$ ways); multiplying these two expressions and then doubling them (because the roles of P_1 and P_2 are different and, thus, distinguishable) gives

$$2\,(k+1)\,(2n-3-k)$$

ways.

iv) There were exactly $k+2$ couples sitting together previous to the n^{th} couple's arrival; P_1 sits between one of the $(k+2)$ couples and P_2 sits between one of the remaining $(k+1)$ couples, giving

$$(k+2)\,(k+1)$$

ways.

Summarizing, for $n > 2$ and $0 \le k \le n$,

$$\begin{aligned} C_{k,n} &= 2(2n-k-1))C_{k-1,n-1} + ((2n-k-2)\,(2n-k-3)+2k)\,C_{k,n-1} \\ &\quad +2(k+1)(2n-k-3)C_{k+1,n-1} + (k+2)(k+1)C_{k+2,n-1}. \end{aligned}$$

See Listing S-3.2 for a program to compute this.

```
function couplesaroundatable(m)
M=zeros(m+1,m+3);        % Keep the pmf in the matrix M.
M(1,2)=1; M(2,1)=2; M(2,3)=4;
for n=3:m
   for k=0:n
      if k==0, a=0;      % a, b, c and d are the values of the four
      else               % predecessors in the recursive formula
       a=M(n-1,k);
      end;
      b=M(n-1,k+1); c=M(n-1,k+2); d=M(n-1,k+3);
      M(n,k+1)=2*(2*n-k-1)*a+((2*(n-1)-k)*(2*(n-1)-k-1)+2*k)*b ...
        +2*(k+1)*(2*(n-1)-k-1)*c+(k+2)*(k+1)*d;
   end;
end; for n=1:m
    M(n,:)=M(n,:)./gamma(2*n); % make it a density!
end;
M(m+1,:)=1./gamma((0:m+2)+1).*exp(-1);
                         % put the Poisson mass function in the last row
M(m:m+1,:)               % show what you have done
E=zeros(m+1,2);          % initialize matrix for the mean values
E(:,1)=M*(0:m+2)';       % calculate the mean values
E(:,2)=(2*(1:(m+1))')./(2*(1:(m+1))'-1);
E(m+1,2)=1               % correct the mean value of the Poisson pmf
```

Program Listing S-3.2: Calculate the probability in the couples-around-a-table problem via recursive use of $C_{k,n}$. It also computes the expected value (see Chapter 4) directly and compares to closed–form solution

d) Event L_k can be written as the sum of $n-k+1$ disjoint events,

$$\Pr(L_k) = \sum_{i=k}^{n} \Pr(F_i),$$

where $F_k = E_k$ and F_{k+j}, $j = 1, \ldots, n-k$, is the event that E_k occurs and exactly j of $C_{k+1}, \ldots C_n$ occur. There are $\binom{n-k}{j}$ ways of picking j of the $C_{k+1}, \ldots C_n$ and each of these possibilities for F_{k+j} is equally likely, and so is equal to $\Pr(E_{k+j})$. Thus,

$$\Pr(L_k) = \sum_{i=k}^{n} \Pr(F_i) = \sum_{i=k}^{n} \binom{n-k}{i-k} \Pr(E_i),$$

which is (3.24). This is easily expressed as (3.25), which can be recursively solved for $\Pr(E_k)$ by using the starting condition

$$\Pr(E_n) = \Pr(L_n),$$

with $\Pr(L_n)$ available from (3.3) with $i = n$, i.e.,

$$\Pr(L_n) = \Pr(C_1 \cdots C_i) = 2^n \frac{(n-1)!}{(2n-1)!}.$$

Finally, as there are $\binom{n}{k}$ possibilities of choosing k specific couples out of n,

$$\Pr(K = k) = \binom{n}{k} \Pr(E_k).$$

e) Because they are disjoint,

$$\Pr(L_k) = \Pr(E_k) + P\left(L_k \cap (C_{k+1} \cup \ldots \cup C_n)\right),$$

so that, from Poincaré's theorem,

$$\begin{aligned}
\Pr(E_k) &= \Pr(L_k) - P((L_k \cap C_{k+1}) \cup \cdots \cup (L_k \cap C_n)) \\
&= \Pr(L_k) - (n-k)\Pr(L_k \cap C_{k+1}) + \binom{n-k}{2}\Pr(L_k \cap C_{k+1} \cap C_{k+2}) - \cdots \\
&= \sum_{i=k}^{n} (-1)^{i-k}\binom{n-k}{n-i}\Pr(L_i) \\
&= \sum_{i=k}^{n} (-1)^{i-k}\binom{n-k}{n-i} 2^i \frac{(2n-i-1)!}{(2n-1)!}.
\end{aligned}$$

using (3.3). As there are $\binom{n}{k}$ ways of picking k couples,

$$\Pr(K = k) = \binom{n}{k}\sum_{i=k}^{n} (-1)^{i-k}\binom{n-k}{n-i} 2^i \frac{(2n-i-1)!}{(2n-1)!}.$$

This agrees with (3.4) because $\binom{i}{k}\binom{n}{i} = \binom{n-k}{n-i}\binom{n}{k}$.

f) We have to show

$$\begin{aligned}
\Pr(E_k) &= \sum_{j=k}^{n} (-1)^{j+k}\binom{n-k}{n-j}\frac{2^j(2n-j-1)!}{(2n-1)!} \\
&=: \sum_{j=k}^{n} (-1)^{j+k}\binom{n-k}{n-j} N_j, \qquad \text{(S-3.2)}
\end{aligned}$$

which, for $k = n$, clearly holds, as the formula simplifies directly to (3.3) with $i = n$. Now assume it holds for $E_{k+1}, E_{k+2}, \ldots, E_n$. Using (3.25), i.e.,

$$\Pr(E_k) = \Pr(L_k) - \sum_{i=k+1}^{n}\binom{n-k}{i-k}\Pr(E_i),$$

we have, using

$$\begin{aligned}
\binom{n-k}{i-k}\binom{n-i}{n-j} &= \frac{(n-k)!}{(i-k)!\,(n-i)!}\frac{(n-i)!}{(n-j)!\,(j-i)!} \\
&= \frac{(n-k)!}{(n-j)!\,(j-k)!}\frac{(j-k)!}{(i-k)!\,(j-i)!} = \binom{n-k}{n-j}\binom{j-k}{i-k}
\end{aligned}$$

and (S-3.2) for $\Pr(E_i)$, $i = k+1, \ldots, n$,

$$\begin{aligned}
& - \sum_{i=k+1}^{n}\binom{n-k}{i-k}\Pr(E_i) \\
&= - \sum_{i=k+1}^{n}\binom{n-k}{i-k}\sum_{j=i}^{n}(-1)^{j+i}\binom{n-i}{n-j}N_j \\
&= - \sum_{j=k+1}^{n}\sum_{i=k+1}^{j}(-1)^{j+i}\binom{n-k}{i-k}\binom{n-i}{n-j}N_j \\
&= - \sum_{j=k+1}^{n}(-1)^{j+k}N_j\sum_{i=k+1}^{j}(-1)^{i-k}\binom{n-k}{n-j}\binom{j-k}{i-k}
\end{aligned}$$

so that, with $m = i - k$ and using the binomial theorem,

$$\begin{aligned}
&= -\sum_{j=k+1}^{n} (-1)^{j+k} \binom{n-k}{n-j} N_j \sum_{m=1}^{j-k} (-1)^m \binom{j-k}{m} \\
&= -\sum_{j=k+1}^{n} (-1)^{j+k} \binom{n-k}{n-j} N_j \left((-1+1)^{j-k} - 1 \right) \\
&= \sum_{j=k+1}^{n} (-1)^{j+k} \binom{n-k}{n-j} N_j
\end{aligned}$$

and, finally, with

$$\Pr(L_k) = N_k = \frac{2^k (2n-k-1)!}{(2n-1)!},$$

$$\Pr(L_k) - \sum_{i=k+1}^{n} \binom{n-k}{i-k} \Pr(E_i) = \sum_{j=k}^{n} (-1)^{j+k} \binom{n-k}{n-j} N_j = \Pr(E_k).$$

Solution to Problem 3.6:

a) Simulation involves randomly permuting a $2n$–length vector with n zeros and n ones, and then checking if element i equals elements $i-1$ and $i+1$, being careful about the two cases arising from the circularity. This is shown in Listing S-3.3. The program was run with 500,000 replications, for three different values of n. A plot of the probabilities for $n = 4$ and $n = 10$ are shown in Figure S-3.2, while the solid line in Figure S-3.3 corresponds to $n = 50$.

From the latter plot, it appears that the probabilities $\Pr(S = s)$ take the shape of the "bell curve". A bit more precisely, (using words and concepts from Chapter 4) the pmf of S approaches a normal pdf. The sample mean and variance are $\mu = 24.24$ and $\sigma^2 = 30.66$, respectively, and a normal pdf with these values is overlaid in Figure S-3.3 as the dashed line. The vertical line is at the mean of 24.24. Clearly, the normal approximation is very good for n this large.

```
function g=surrounded(n)
sim=500000; trapped=zeros(sim,1); v=[zeros(1,n) ones(1,n)];
for s=1:sim
  if mod(s,1000)==0, [n,s], end
  tab=v(randperm(2*n));  mcnt=0;
  if (tab(end)==tab(1)) & (tab(1)==tab(2)), mcnt=mcnt+1; end
  if (tab(end-1)==tab(end)) & (tab(end)==tab(1)), mcnt=mcnt+1; end
  for i=2:(2*n-1)
    if (tab(i-1)==tab(i)) & (tab(i)==tab(i+1)), mcnt=mcnt+1; end
  end
  trapped(s)=mcnt;
end
g=tabulate(trapped+1);
% plot(g(:,1),g(:,3)/100)
% mean(trapped), std(trapped)
```

Program Listing S-3.3: Simulates probability of getting "trapped" between two people of the same sex when sitting at a round table with n couples.

b) Program (S-3.4) computes both the exact values of (3.27) and simulated ones. Running the code

```
[f,fsim]=holst(5,50000);
plot(0:5,f,'r-',0:length(fsim)-1,fsim,'g--')
```

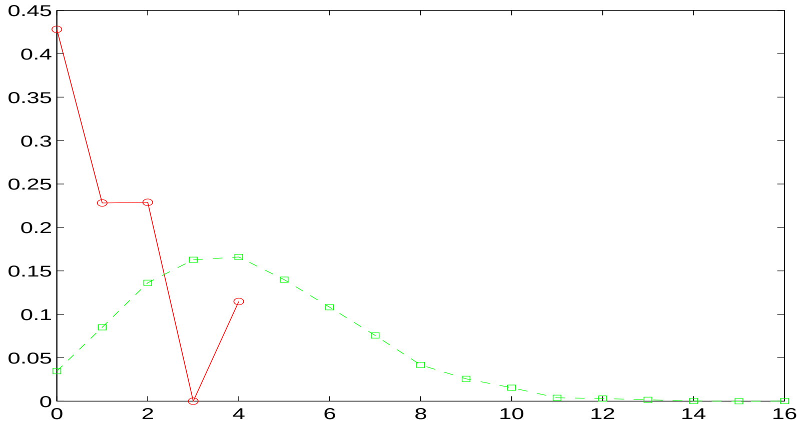


Figure S-3.2: Simulated probability of getting "trapped" between two people of the same sex when sitting at a round table with n couples, based on 500,000 replications. Solid is for $n = 4$, and dashed is for $n = 10$.

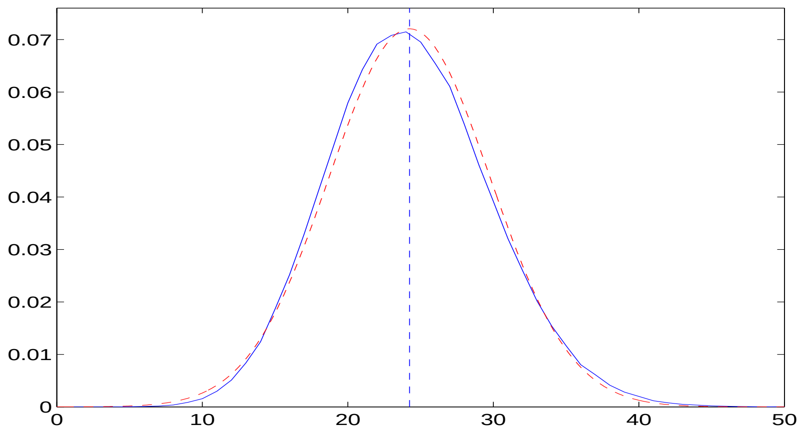


Figure S-3.3: Same as in Figure S-3.2 but for $n = 50$ (solid) and the normal pdf with mean and variance chosen to the be the sample mean and sample variance, respectively, from the 500,000 replications.

and

```
[f,fsim]=holst(50,50000);
plot(0:50,f,'r-',0:length(fsim)-1,fsim,'g--'), axis([0 8 0 0.3])
```

produces graphs showing the exact and simulated (with $50,000$ replications) probabilities for $n = 5$ and $n = 50$ which are virtually indistinguishable and are not shown.

Figure S-3.4 shows both (3.4) and (3.27) for $n = 5$ and $n = 50$. In both cases, the "density" of (3.27) is, approximately speaking, shifted to the right, compared to (3.4). This makes sense: One might think that, when wives are already placed in every other seat, the chances are higher for more matches than when all $2n$ people are randomly dispersed around the table. But, as the probabilities sum to one, they cannot be larger for *each* s, i.e., the number of couples sitting together. Instead, under (3.27), i.e., when wives occupy every other seat, the probabilities are larger *only when s is large*. This is because it is even *more* unlikely that a large number of couples would be sitting together when all $2n$ people are randomly placed.

Solution to Problem 3.7:

a) The probability that at least 10 people order beef is

$$(0.5)^{12}\left\{\binom{12}{10}+\binom{12}{11}+\binom{12}{12}\right\}=\frac{79}{2^{12}},$$

which, by symmetry, is also the probability that at least 10 people order fish, so that we have

```
function [f,fsim]=holst(n,sim)
f=zeros(n+1,1);
for w=0:n
  r=w:n;
  temp = (-1).^(r-w) .* c(r,w) ./ gamma(r+1) .* (2*n) ./ (2*n-r);
  f(w+1)= sum(  temp  .* c(2*n-r,r) ./ c(n,r) );
end
if nargin<2, sim=0; end
if sim>0
  numpair=zeros(sim,1); tab=zeros(2*n,1);
  for i=1:2:(2*n-1), tab(i)=(i-1)/2+1; end
  for s=1:sim
    men=randperm(n);
    for i=2:2:2*n, tab(i)=men(i/2); end
    mcnt = (tab(1)==tab(end));
    for i=1:(2*n-1)
      if tab(i)==tab(i+1), mcnt=mcnt+1; end
    end
    numpair(s)=mcnt;
  end
  tt=tabulate(numpair+1);
  fsim=tt(:,3)/100;
end
```

Program Listing S-3.4: Computes (3.27) if `sim`= 0 and simulates it otherwise

b) $2 \cdot 79 \cdot 2^{-12} \approx 3.8574 \times 10^{-2}$.

c) The probability that at least 10 people order beef should be rather small, although it is still possible. The probability that at least 10 people order fish is the same as the probability that either 0,1 or 2 people order beef. We have

$$\sum_{i=10}^{12} \binom{12}{i} 0.1^i 0.9^{12-i} + \sum_{i=0}^{2} \binom{12}{i} 0.1^i 0.9^{12-i} \approx 5.455 \times 10^{-9} + 0.88913.$$

Solution to Problem 3.8: Although this appears to be the same as the question in Problem 2.19 ("Your new neighbors have 2 children, and at least one is a boy. What is the probability that the other is also a boy?"), the answer is 1/2. To see this formally, denoting B for boy, G for girl, event $\{BG\}$ stands for "first child boy, second child girl", etc., and b is the event that a boy is seen in the

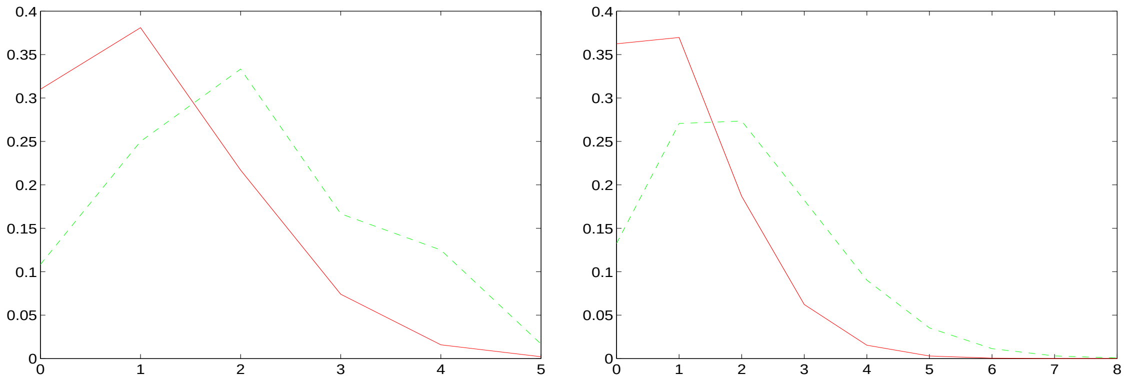


Figure S-3.4: Comparison of (3.4) (solid line) and (3.27) (dashed) for $n = 5$ (left) and $n = 50$) (right) garden, then, from Bayes' rule (and implicitly conditional on the knowledge that the neighbor has exactly 2 children and both are equally likely regardless of sex to play in the garden), $\Pr(BB \mid b)$ is given by

$$\frac{\Pr(b \mid BB)\Pr(BB)}{\Pr(b \mid BB)\Pr(BB)+\Pr(b \mid BG)\Pr(BG)+\Pr(b \mid GB)\Pr(GB)+\Pr(b \mid GG)\Pr(GG)}$$

or

$$\frac{1\cdot(1/4)}{1\cdot(1/4)+(1/2)(1/4)+(1/2)(1/4)+0\cdot(1/4)}=\frac{1}{2}.$$

To see intuitively how this differs from Problem 2.19, consider "simulating" to get the answer. With equal probability 1/4 we get a neighbor with either BB, BG, GB or GG. Next to each "realization" we place a one to indicate that a boy is playing in the garden, or a zero if not. This occurs with probability zero for GG, probability 1/2 for GB and BG, and one for BB. The following four realizations are a representative set of occurrences (because BG gets a one and GB gets a zero, which reflects the average probability of 1/2 for BG and GB):

$$BB:1,\quad GG:0,\quad BG:1,\quad GB:0$$

Because we condition on observing a boy in the garden, our sample consists only of those pairs followed by a 1, i.e., BB, BG. Thus, from our restricted sample, the probability of having 2 boys is 1/2.

Solution to Problem 3.9:

a) $\Pr(S_1S_2)=\Pr(S_2 \mid S_1)\Pr(S_1)=\dfrac{s+c}{s+c+r}\dfrac{s}{s+r}$.

b) Directly, $\Pr(S_2 \mid R_1)=\dfrac{s}{s+c+r}$ and, from Bayes' rule,

$$\begin{aligned}\Pr(S_1 \mid R_2) &= \frac{\Pr(R_2 \mid S_1)\Pr(S_1)}{\Pr(R_2 \mid S_1)\Pr(S_1)+\Pr(R_2 \mid R_1)\Pr(R_1)}\\ &= \frac{\frac{r}{s+c+r}\cdot\frac{s}{s+r}}{\frac{r}{s+c+r}\cdot\frac{s}{s+r}+\frac{r+c}{s+c+r}\cdot\frac{r}{s+r}}\\ &= \frac{s}{s+c+r}\end{aligned}$$

so that $\Pr(S_2 \mid R_1)=\Pr(S_1 \mid R_2)$.

c) For the former,

$$\begin{aligned}\Pr(S_3 \mid R_1) &= \Pr(S_3 \mid R_1,R_2)\Pr(R_2 \mid R_1)+\Pr(S_3 \mid R_1,S_2)\Pr(S_2 \mid R_1)\\ &= \frac{s}{s+r+2c}\frac{r+c}{s+r+c}+\frac{s+c}{r+s+2c}\frac{s}{r+s+c}=\frac{s}{s+c+r}.\end{aligned}$$

For the latter,

$$\Pr(S_1 \mid R_3)=\frac{\Pr(R_3 \mid S_1)\Pr(S_1)}{\Pr(R_3 \mid S_1)\Pr(S_1)+\Pr(R_3 \mid R_1)\Pr(R_1)},$$

but

$$\begin{aligned}\Pr(R_3 \mid S_1) &= \Pr(R_3 \mid S_1,S_2)\Pr(S_2 \mid S_1)+\Pr(R_3 \mid S_1,R_2)\Pr(R_2 \mid S_1)\\ &= \frac{r}{r+s+2c}\frac{s+c}{s+c+r}+\frac{r+c}{r+s+2c}\frac{r}{s+r+c}=\frac{r}{s+c+r}\end{aligned}$$

and $\Pr(R_3 \mid R_1)=1-\Pr(S_3 \mid R_1)=\dfrac{r+c}{s+c+r}$ so that

$$\Pr(S_1 \mid R_3)=\frac{\frac{r}{s+c+r}\frac{s}{s+r}}{\frac{r}{s+c+r}\frac{s}{s+r}+\frac{r+c}{s+c+r}\frac{r}{s+r}}=\frac{s}{s+c+r},$$

showing that $\Pr(S_3 \mid R_1)=\Pr(S_1 \mid R_3)$.

Solution to Problem 3.10: For $1 \leq n \leq 13$, given that the $n-1$ cards 2 through n are hearts, there is a $(13-(n-1))/(52-(n-1))$ chance that any other card (card 1 in particular) will be a heart, i.e.,

$$P_n = \begin{cases} \frac{14-n}{53-n}, & 1 \leq n \leq 13, \\ 0, & \text{otherwise.} \end{cases}$$

More formally,

$$P_n = \frac{\frac{13}{52} \cdots \frac{13-n+1}{52-n+1}}{\frac{13}{52} \cdots \frac{13-n+1}{52-n+1} + \frac{39}{52}\left(\frac{13}{51} \cdots \frac{13-n+2}{51-n+2}\right)},$$

but the denominators of all fractions are the same, so that

$$P_n = \frac{\frac{13!}{(13-n)!}}{\frac{13!}{(13-n)!} + 39\left(\frac{13!}{(13-n)!(13-n+1)}\right)} = \frac{1}{1+\frac{39}{13-n+1}} = \frac{14-n}{53-n}.$$

Solution to Problem 3.11: Define B_i to be the event the object is in box i, $i = 1, 2, 3$ and F is the event that the first search of box 3 fails. Then Bayes' rule gives

$$\Pr(B_i \mid F) = \frac{\Pr(F \mid B_i)\Pr(B_i)}{\sum_{i=1}^{3}\Pr(F \mid B_i)\Pr(B_i)},$$

with $\Pr(F \mid B_1) = 1$, $\Pr(F \mid B_2) = 1$ and $\Pr(F \mid B_3) = 0.3$. $\Pr(B_2 \mid F) = 0.462$ is the highest of the 3 and is, thus, the answer.

Notice that the denominator is $\Pr(F) = 0.2 + 0.3 + (0.3)(0.5)$, while the numerator is one of these three components for each i. The denominator is just a "correction factor" such that the probabilities add to one. As such, it makes sense just to compare values 0.2, 0.3 and $(0.3)(0.5) = 0.15$, showing that box 2 should be searched. Thus, the proposed "naive solution" was correct. If one wants the conditional probabilities, then these values have to be divided by $\Pr(F) = 0.65$ to give respective box probabilities 0.308, 0.462 and 0.231.

Solution to Problem 3.12: This is (3.18) with $n = 4$, $m = 4$ and $p = 0.6$, so that

$$Pr(A \text{ wins}) = \sum_{k=4}^{7}\binom{7}{k}0.6^k 0.4^{7-k} = 0.710208.$$

Solution to Problem 3.13: Substituting into (3.20),

$$\begin{aligned}
\Pr(A_{1,m}) &= p\Pr(A_{0,m}) + (1-p)\Pr(A_{1,m-1}) \\
&= p + (1-p)\left[p\Pr(A_{0,m-1}) + (1-p)\Pr(A_{1,m-2})\right] \\
&= p + p(1-p) + (1-p)^2\Pr(A_{1,m-2}) \\
&= p + p(1-p) + (1-p)^2\left[p\Pr(A_{0,m-2}) + (1-p)\Pr(A_{1,m-3})\right] \\
&= p + p(1-p) + p(1-p)^2 + (1-p)^3\Pr(A_{1,m-3}) \\
&\;\;\vdots \\
&= p + p(1-p) + p(1-p)^2 + \cdots + p(1-p)^{m-1} \\
&= p\frac{1-(1-p)^m}{1-(1-p)} = 1-(1-p)^m.
\end{aligned}$$

This agrees with (3.16), i.e., with $n = 1$,

$$\Pr(A_{1,m}) = \sum_{i=0}^{m-1}\binom{1+i-1}{i}p^1(1-p)^i = p\sum_{i=0}^{m-1}(1-p)^i = 1-(1-p)^m.$$

Solution to Problem 3.14:

a) See Program S-3.5 and note the use of recursive function calls.

b) Use of (3.16) is slightly more advantageous because p^n can be factored out. See Program S-3.6, which makes judicious use of the vector ability of Matlab.

c)

i) See Program S-3.7, which also demonstrates use of the `while` construction.

ii) See Program S-3.8. Note the use of the `nargin` function, as well as the `mesh` 3D graphics function. The graph from calling `plotgmin(0.65:0.01:0.9)` is shown in Figure S-3.5.

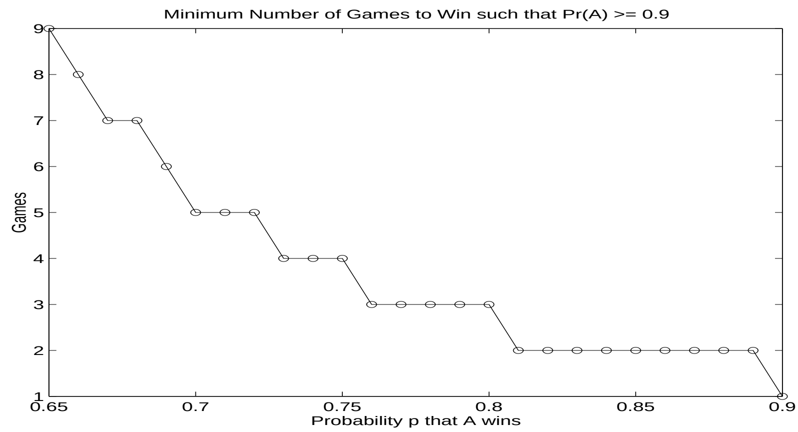


Figure S-3.5: Output from `plotgmin(0.65:0.01:0.9)`

Solution to Problem 3.15: With $P_n := \Pr(A_{n,n})$ and $r := p(1-p)$, (3.16) and (1.4) give

$$\begin{aligned}
\frac{1}{p}P_{n+1} &= \sum_{i=0}^{n}\binom{n+i}{i}p^n(1-p)^i \\
&= \sum_{i=0}^{n-1}\binom{n+i-1}{i}p^n(1-p)^i + \binom{2n-1}{n}r^n + \sum_{i=1}^{n}\binom{n+i-1}{i-1}p^n(1-p)^i \\
&= P_n + \frac{1}{2}\binom{2n}{n}r^n + \sum_{i=0}^{n-1}\binom{n+i}{i}p^n(1-p)^{i+1} \\
&= P_n + \frac{1}{2}\binom{2n}{n}r^n + \frac{1-p}{p}P_{n+1} - (1-p)\binom{2n}{n}r^n \\
&= P_n + \frac{1-p}{p}P_{n+1} + \left\{\frac{1}{2} - (1-p)\right\}\binom{2n}{n}r^n.
\end{aligned}$$

This yields the recursive formula

$$\left\{\frac{1}{p} - \frac{1-p}{p}\right\}P_{n+1} = P_{n+1} = P_n + \left(p - \frac{1}{2}\right)\binom{2n}{n}r^n$$

with $P_1 = p$; solving gives (3.28).

Solution to Problem 3.16: Clearly, T has be to be greater than both n and m, before which neither A nor B can win. Once $T = t \geq n$, then A can win, and this occurs with probability $\binom{t-1}{n-1}p^n(1-p)^{t-n}$. Likewise, once $t \geq m$, B can win. If t is greater than both n and m, then either A or B could win, thus contributing to the probability that $T = t$. Putting this together,

To compute $\Pr(T \mid A_{n,m})$, use Bayes' rule (3.12) to get

$$\Pr(T = t \mid A_{n,m}) = \frac{\binom{t-1}{n-1}p^n(1-p)^{t-n}\,\mathbb{I}_{\{n,n+1,\ldots\}}(t)}{\Pr(A_{n,m})}.$$

```
function P=fermat(n,m,p)
if n==0
  P=1;
elseif m==0
  P=0;
else
  P=p * fermat(n-1,m,p) + (1-p) * fermat(n,m-1,p);
end
```

Program Listing S-3.5: Fermat's solution to the problem of the points

```
function P = PoP(n,m,p)
i=0:(m-1); P=p^n * sum( c(n+i-1,i) .* (1-p).^i );
```

Program Listing S-3.6: Closed form solution to the problem of the points

```
function g=gmin(a,p)
g=0; done=0;
while not(done)
  g=g+1;
  done=(PoP(g,g,p) >= a);
end
```

Program Listing S-3.7: Smallest g such that $\Pr(A \mid g, g, p) \geq \alpha$

```
function plotgmin(pvec,avec)
if nargin < 2, avec=0.90; end
plen=length(pvec); alen=length(avec); g=zeros(alen,plen);
for aloop=1:alen
  a=avec(aloop);
  for ploop=1:plen
    p=pvec(ploop);  g(aloop,ploop)=gmin(a,p);
  end
end
if alen==1
  plot(pvec,g,'o',pvec,g,'r-')
  xlabel ('Probability p that A wins'), ylabel('Games')
  title (['Min # of Games to Win such that Pr(A) >= ' num2str(avec)])
else
  mesh(pvec,avec,g)
  title ('Min # of Games to Win such that Pr(A) > \alpha')
  xlabel ('Probability p that A wins')
  ylabel ('Tolerance \alpha'), view(-25,15)
end
```

Program Listing S-3.8: Plot g for various p (pvec) and α (avec)

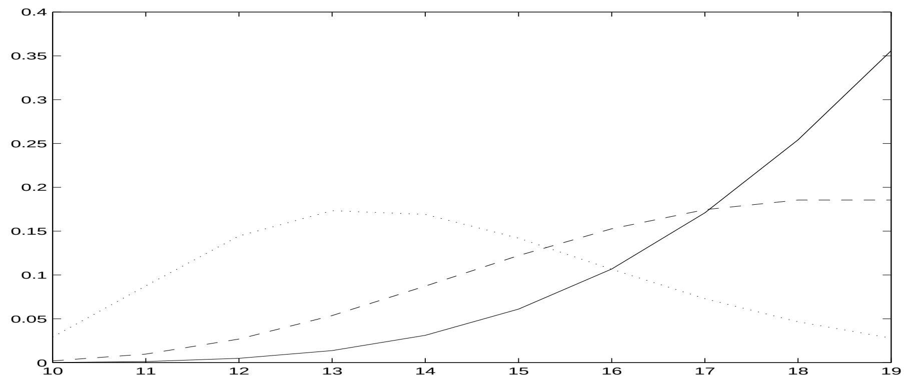


Figure S-3.6: Posterior $\Pr(T = t \mid A_{10,10})$ versus t for $p = 0.3$ (solid) 0.5 (dashed) and 0.7 (dotted)

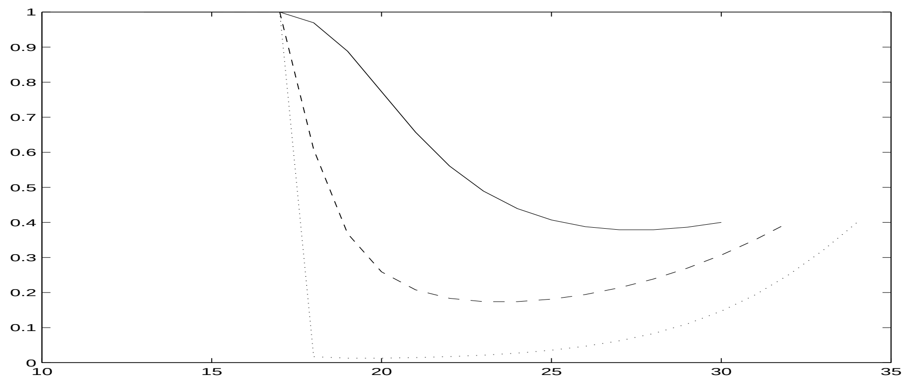


Figure S-3.7: Posterior $\Pr(A_{n,18} \mid T = t)$ versus t for $p = 0.4$ and $n = 13$ (solid) 15 (dashed) and 17 (dotted)

Likewise, the conditional probability that A wins, given that $T = t$ rounds were played, is

$$\Pr(A_{n,m} \mid T = t) = \frac{\binom{t-1}{n-1} p^n (1-p)^{t-n} \mathbb{I}_{\{n,n+1,\ldots\}}(t)}{\Pr(T = t)}.$$

Clearly, if $p = 0.5$, then $\Pr(A_{n,n} \mid T = t) = 0.5$ for all $n \leq T \leq 2n - 1$.

Figure S-3.6 plots $\Pr(T = t \mid A_{10,10})$ versus t for $p = 0.3, 0.5$ and 0.7; while Figure S-3.7 plots $\Pr(A_{n,18} \mid T = t)$ versus t for $p = 0.4$ and $n = 13, 15$ and 17.

Solution to Problem 3.17: To show that the sum in (3.29) is absolutely convergent for $p < 1/2$, let $a_n = \binom{2n}{n} [p(1-p)]^n$ so that, by l'Hôpital's rule,

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n\to\infty} \left[\frac{(2n+2)(2n+1)}{(n+1)^2} \right] p(1-p) = 4p(1-p) < 1,$$

because $p(1-p) < \frac{1}{4}$ for $p < \frac{1}{2}$. The latter follows directly from $\left(p - \frac{1}{2}\right)^2 > 0$ for $p \neq \frac{1}{2}$.

From (3.28), $0 = p + \left(p - \frac{1}{2}\right) S$, which implies

$$S = \frac{2p}{1-2p} = 2p \sum_{i=0}^{\infty} 2^i p^i = \sum_{i=1}^{\infty} 2^i p^i,$$

or (3.30), i.e.,

$$\sum_{i=1}^{\infty} \binom{2i}{i} [p(1-p)]^i = \sum_{i=1}^{\infty} 2^i p^i.$$

Using the binomial theorem, the lhs may also be written as

$$\sum_{i=1}^{\infty}\binom{2i}{i}\left[p-p^2\right]^i=\sum_{i=1}^{\infty}\binom{2i}{i}\sum_{j=0}^{i}\binom{i}{j}(-1)^j\,p^{2j}p^{i-j}=\sum_{i=1}^{\infty}\sum_{j=0}^{i}\binom{2i}{i}\binom{i}{j}(-1)^j\,p^{i+j}.$$

Clearly, $y := i+j$ goes from 1 to ∞; collecting those terms yields

$$\sum_{i=1}^{\infty}\sum_{j=0}^{i}\binom{2i}{i}\binom{i}{j}(-1)^j\,p^{i+j}=\sum_{y=1}^{\infty}\sum_{i=1}^{y}\binom{2i}{i}\binom{i}{y-i}(-1)^{y-i}\,p^y.$$

Comparing to the rhs of (3.30), this gives the identity (3.31), i.e.,

$$2^y=\sum_{i=1}^{y}\binom{2i}{i}\binom{i}{y-i}(-1)^{i+y}=\sum_{i=\lceil y/2\rceil}^{y}\binom{2i}{y}\binom{y}{i}(-1)^{i+y},$$

where $\lceil x\rceil=\ \mathrm{ceil}\,(x)$ and the latter term follows from $\binom{2i}{i}\binom{i}{y-i}=\binom{2i}{y}\binom{y}{i}$ and the fact that $i\geq y-i\Rightarrow i\geq y/2$.

Solution to Problem 3.18: We have, with $(1-2p)^2=1-4p\,(1-p)$ and $m=n-1$,

$$\begin{aligned}
g'(p)\,(1-2p) &= \sum_{n=0}^{\infty}\binom{2n}{n}n\,[p\,(1-p)]^{n-1}\,(1-2p)^2\\
&= \sum_{n=0}^{\infty}\binom{2n}{n}n\,[p\,(1-p)]^{n-1}-\sum_{n=0}^{\infty}n\binom{2n}{n}[p\,(1-p)]^{n-1}\,4p\,(1-p)\\
&= \sum_{n=1}^{\infty}\binom{2n}{n}n\,[p\,(1-p)]^{n-1}-\sum_{n=0}^{\infty}4n\binom{2n}{n}[p\,(1-p)]^{n}\\
&= \sum_{m=0}^{\infty}\binom{2m+2}{m+1}(m+1)\,[p\,(1-p)]^{m}-\sum_{m=0}^{\infty}4m\binom{2m}{m}[p\,(1-p)]^{m}\\
&= \sum_{m=0}^{\infty}\left\{\frac{(2m+2)\,(2m+1)}{m+1}-4m\right\}\binom{2m}{m}[p\,(1-p)]^{m}\\
&= 2\sum_{m=0}^{\infty}\binom{2m}{m}[p\,(1-p)]^{m}=2g\,(p)\,,
\end{aligned}$$

so that $g'(p)\,(1-2p)=2g\,(p)$ and the result follows.

Solution to Problem 3.19:

a) Note that, if A wins the first round played, then (because of independence of trials) the game can be viewed as "starting over" but such that now A has $i+1$ dollars and B has $T-i-1$ dollars. Thus, $\Pr_i(A\mid W)=\Pr_{i+1}(A)=s_{i+1}$. Using the law of total probability (3.10), it follows that

$$\begin{aligned}
s_i=\Pr_i(A) &= \Pr_i(A\mid W)\Pr(W)+\Pr_i\left(A\mid\bar{W}\right)\Pr\left(\bar{W}\right)\\
&= s_{i+1}\,p+s_{i-1}\,q,
\end{aligned}$$

i.e.,

$$s_i=ps_{i+1}+qs_{i-1},\qquad 1\leq i\leq T,$$

or, as $s_i=ps_i+qs_i$,

$$qs_i-qs_{i-1}=ps_{i+1}-ps_i.$$

With $r=q/p$ and $d_i=s_{i+1}-s_i$, this yields $d_i=rd_{i-1}$ or $d_i=r^id_0$.

Conditioning on $i = 0$, we see that $s_0 = 1$. Similarly, $s_T = 0$. Then

$$1 = s_0 - s_T = -\sum_{i=0}^{T-1} d_i = -d_0 \sum_{i=0}^{T-1} r^i = -d_0 \frac{1 - r^T}{1 - r}$$

so that

$$d_0 = -\frac{1 - r}{1 - r^T}.$$

Similarly,

$$s_j = s_j - 0 = s_j - s_T = -d_0 \sum_{i=j}^{T-1} r^i = -d_0 \frac{r^j - r^T}{1 - r},$$

so that

$$s_j = (-1)\frac{1 - r}{1 - r^T}(-1)\frac{r^j - r^T}{1 - r} = \frac{r^j - r^T}{1 - r^T}, \qquad 0 \le j \le T.$$

b) We have

$$\lim_{p \to \frac{1}{2}} \frac{1 - \left(\frac{1-p}{p}\right)^i}{1 - \left(\frac{1-p}{p}\right)^T}$$

is indeterminant, l'Hôpital's rule implies

$$\lim_{p \to \frac{1}{2}} \frac{\frac{d}{dp}\left(1 - \left(\frac{1-p}{p}\right)^i\right)}{\frac{d}{dp}\left(1 - \left(\frac{1-p}{p}\right)^T\right)} = \lim_{p \to \frac{1}{2}} \frac{\frac{i}{p(1-p)}\left(\frac{1}{p}(1-p)\right)^i}{\frac{T}{p(1-p)}\left(\frac{1}{p}(1-p)\right)^T} = \frac{4i}{4T} = \frac{i}{T}.$$

Solution to Problem 3.20: See Program S-3.9 and output shown in Figure S-3.8.

```
function disease
subplot(2,1,1), a=0:80; plot(a,postd(a,10)), xlabel('age')
title ('Pr(D|+) For c=10 Cigarettes')

subplot(2,1,2), c=0:40; plot(c,postd(45,c))
xlabel('Cigarettes per Day'), title ('Pr(D|+) For age=45 Years')
orient tall

function p=postd(a,c)
b=a/20-3; d=(c+1)*(tanh(b)+1)/100; p=0.95*d ./ (0.95*d + 0.02*(1-d));
```

Program Listing S-3.9: $\Pr(D \mid +)$ for $0 \le a \le 80$, $c = 10$ and $\Pr(D \mid +)$ for $a = 45$, $0 \le c \le 40$

Solution to Problem 3.21: The function f_n is given by $c_1\lambda_1^n + c_2\lambda_2^n$, where $\lambda_{1,2}$ are the solutions of the characteristic equation $\lambda^2 - (1-p)\lambda - p(1-p) = 0$, i.e.,

$$2\lambda_{1,2} = 1 - p \pm \sqrt{(1 + 2p - 3p^2)}.$$

Using the starting values f_1 and f_2, the solution to the two equations

$$c_1\lambda_1 + c_2\lambda_2 = 0, \quad c_1\lambda_1^2 + c_2\lambda_2^2 = p^2$$

gives the values of coefficients c_1 and c_2, i.e., with $w = \sqrt{(1 + 2p - 3p^2)}$,

$$c_1 = \frac{p^2}{\lambda_1(\lambda_1 - \lambda_2)} = \frac{-2p^2}{(p - w - 1)\,w}, \quad c_2 = \frac{p^2}{\lambda_2(\lambda_2 - \lambda_1)} = \frac{2p^2}{(p + w - 1)\,w}.$$

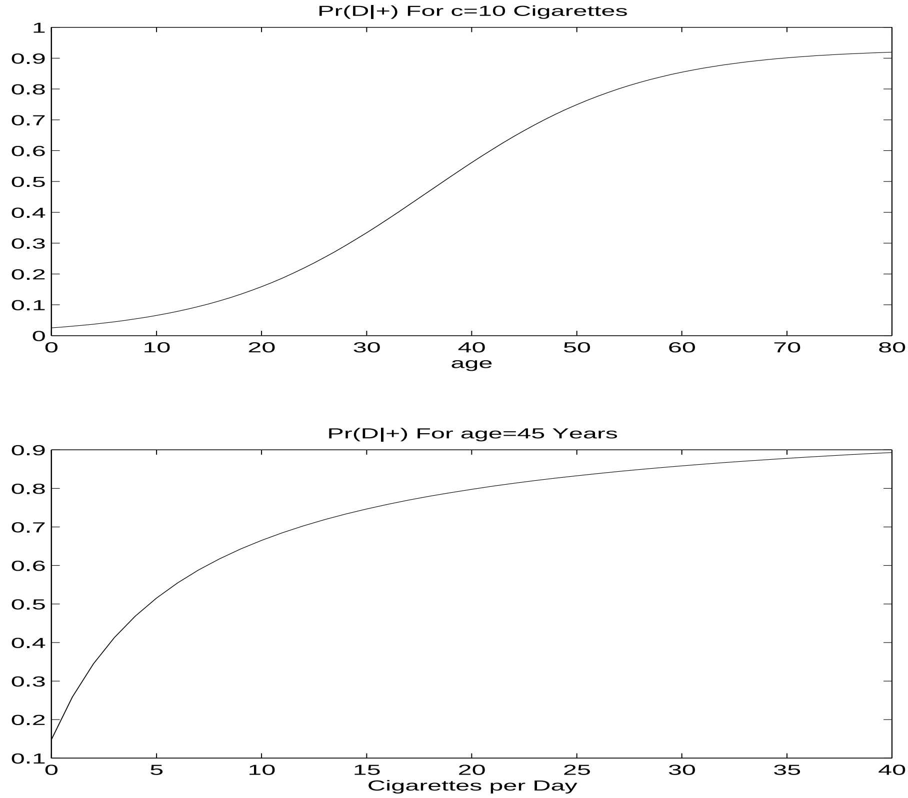


Figure S-3.8: Output from `disease`

That is,

$$\begin{aligned} f_n &= c_1\lambda_1^n + c_2\lambda_2^n \\ &= 2p^2w^{-1}\left(\frac{\lambda_2^n}{p+w-1} - \frac{\lambda_1^n}{p-w-1}\right) \\ &= 2p^2w^{-1}\left(\frac{(1-p-w)^n\,2^{-n}}{p+w-1} - \frac{(1-p+w)^n\,2^{-n}}{p-w-1}\right) \\ &= 2^{1-n}p^2w^{-1}\left((1-p+w)^{n-1} - (1-p-w)^{n-1}\right). \end{aligned}$$

For low values of n, this can be simplified; several are shown in the following table.

n	f_n
1	0
2	p^2
3	$p^2(1-p)$
4	$p^2(1-p)$
5	$p^2(p+1)(1-p)^2$

.

To verify that $\sum_{i=1}^{\infty} f_n = 1$,

$$\begin{aligned} \sum_{i=1}^{\infty} f_i &= p^2w^{-1}\sum_{i=1}^{\infty} 2^{1-i}\left((1-p+w)^{i-1} - (1-p-w)^{i-1}\right) \\ &= p^2w^{-1}\sum_{i=2}^{\infty}\left(\frac{1-p+w}{2}\right)^{i-1} - p^2w^{-1}\sum_{i=2}^{\infty}\left(\frac{1-p-w}{2}\right)^{i-1} \\ &= p^2w^{-1}\left(\frac{1-p+w}{1+p-w}\right) - p^2w^{-1}\left(\frac{1-p-w}{1+p+w}\right) \\ &= 1, \end{aligned}$$

after some simplification using Maple. For the expected value, using Maple gives

$$\begin{aligned}
\sum_{i=1}^{\infty} i f_i &= p^2 w^{-1} \sum_{i=2}^{\infty} i \left(\frac{1-p+w}{2}\right)^{i-1} - p^2 w^{-1} \sum_{i=2}^{\infty} i \left(\frac{1-p-w}{2}\right)^{i-1} \\
&= p^2 w^{-1} \left(\frac{(1-p+w)(3+p-w)}{(1+p-w)^2}\right) - p^2 w^{-1} \left(\frac{(3+p+w)(1-p-w)}{(1+p+w)^2}\right) \\
&= \frac{p+1}{p^2},
\end{aligned}$$

as was to be shown.

Solutions to Chapter 4: Univariate Random Variables

Solution to Problem 4.1: $\Pr(X > 1) = 1 - \Pr(X = 1) = 0.7$ and

$$\Pr(X = 2 \mid X > 1) = \frac{\Pr(X = 2 \cap X > 1)}{\Pr(X > 1)} = \frac{\Pr(X = 2)}{\Pr(X > 1)} \Rightarrow \Pr(X = 2) = 0.56,$$

so that

$$\Pr(X > 2) = 1 - \Pr(X = 1) - \Pr(X = 2) = 1 - 0.3 - 0.56 = 0.14.$$

Solution to Problem 4.2:

$$\Pr(X = 0) + \Pr(X = 1) = \frac{1}{2} \Rightarrow p + pq = \frac{1}{2} \Rightarrow p^2 - 2p + \frac{1}{2} = 0 \Rightarrow p = \frac{2 - \sqrt{2}}{2}.$$

Because $X \sim \text{Geo}(p)$, it follows that

$$\mathbb{E}[X] = \frac{1-p}{p} = \frac{1}{\sqrt{2} - 1} \approx 2.4142.$$

Solution to Problem 4.3: For density (4.16) with $q = 1 - p$,

$$\mathbb{E}[X] = \sum_{x=0}^{\infty} x f_X(x) = p \sum_{x=0}^{\infty} x q^x =: pS_1,$$

where

$$\begin{aligned} S_1 &= q + 2q^2 + 3q^3 + \cdots \\ qS_1 &= \qquad q^2 + 2q^3 + 3q^4 + \cdots \\ S_1 - qS_1 &= q + q^2 + q^3 + \cdots = \frac{q}{1-q} \end{aligned}$$

so that

$$S_1 = \frac{q}{(1-q)^2}, \qquad \mathbb{E}[X] = pS_1 = \frac{1-p}{p}.$$

For $\mathbb{E}\left[X^2\right]$ using density (4.18),

$$\mathbb{E}\left[X^2\right] = \sum_{x=1}^{\infty} x^2 f_X(x) = p \sum_{x=1}^{\infty} x^2 q^{x-1} = p \sum_{x=1}^{\infty} x^2 q^{x-1} =: pS_2,$$

where

$$\begin{aligned} S_2 &= 1 + 4q + 9q^2 + 16q^3 + \cdots + x^2 q^{x-1} + \cdots \\ qS_2 &= \qquad q + 4q^2 + \ 9q^3 + \cdots + (x-1)^2 q^x + \cdots \\ S_2 - qS_2 &= 1 + 3q + 5q^2 + 7q^3 + \cdots + (2x-1) q^{x-1} + \cdots \\ &= \sum_{i=0}^{\infty} (2i+1) q^i = 2S_1 + \sum_{i=0}^{\infty} q^i = 2\frac{q}{(1-q)^2} + \frac{1}{1-q}, \end{aligned}$$

so that

$$S_2 = 2\frac{q}{(1-q)^3} + \frac{1}{(1-q)^2} = \frac{1+q}{(1-q)^3}, \qquad \mathbb{E}\left[X^2\right] = \frac{2-p}{p^2}$$

and

$$\mathbb{V}(X) = \frac{2-p}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{1-p}{p^2},$$

which holds for both density (4.16) and (4.18) because $\mathbb{V}(X-1) = \mathbb{V}(X)$.

Solution to Problem 4.4: From the definition of expected value and the binomial theorem,

$$\begin{aligned} \mathbb{E}[X] &= \sum_{x=0}^{n} x\, f_X(x) = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=1}^{n} \frac{x\, n!}{(n-x)!\, x!} p^x (1-p)^{n-x} \\ &= np \sum_{x=1}^{n} \frac{(n-1)!}{(n-x)!\,(x-1)!} p^{x-1} (1-p)^{n-x} = np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} \\ &= np\left[p + (1-p)\right]^{n-1} = np, \end{aligned}$$

where $k = i - 1$.

Solution to Problem 4.5: Let $Y = \sigma X$. Then, with $\mu = \mathbb{E}[X]$, $\mathbb{E}[Y] = \sigma\mu$ and

$$\frac{\mu_4(Y)}{\mu_2^2(Y)} = \frac{\mathbb{E}\left[(Y - \mathbb{E}[Y])^4\right]}{\left(\mathbb{E}\left[(Y - \mathbb{E}[Y])^2\right]\right)^2} = \frac{\mathbb{E}\left[(\sigma X - \sigma\mu)^4\right]}{\left(\mathbb{E}\left[(\sigma X - \sigma\mu)^2\right]\right)^2} = \frac{\sigma^4 \mathbb{E}\left[(X-\mu)^4\right]}{\sigma^4\left(\mathbb{E}\left[(X-\mu)^2\right]\right)^2} = \frac{\mu_4(X)}{\mu_2^2(X)}.$$

Solution to Problem 4.6:

a) On the first toss, the probability of no sixes is $(5/6)^d$, so that

$$p = 1 - (5/6)^d$$

is the probability of at least one six. The number of tosses (of all d dice) until at least one six occurs is clearly geometrically distributed, so the expected number of rolls until at least one six occurs is $1/p$ or

$$\frac{1}{p} = \frac{1}{1-(5/6)^d} = \frac{6^d}{6^d - 5^d}.$$

On the toss which produced at least one six, the probability of getting exactly one six is

$$\Pr(\text{one six} \mid \text{at least one six}) = \frac{\Pr(\text{one six})}{p} = \frac{1}{p}\binom{d}{1}\frac{1}{6}\left(\frac{5}{6}\right)^{d-1} = d\frac{5^{d-1}}{6^d - 5^d}.$$

Similarly, on the toss which produced at least one six, the probability of getting exactly k sixes is

$$\frac{1}{p}\binom{d}{k}\left(\frac{1}{6}\right)^k\left(\frac{5}{6}\right)^{d-k} = \binom{d}{k}\frac{5^{d-k}}{6^d - 5^d}.$$

The critical step is now: If, say, T initial tosses were required to produce at least one six, and k sixes occurred on that trial, then the expected total number of tosses is $T + \mathbb{E}_{d-k}[N]$. That is, if k is fixed, we expect

$$\frac{1}{p} + \mathbb{E}_{d-k}[N]$$

tosses. Taking the fact that k is random into account and that events $\{k=1\}, \ldots, \{k=d\}$ are disjoint gives (with $j = d - k$ and reversing the order of summation),

$$\begin{aligned} \mathbb{E}_d[N] &= \frac{1}{p} + \sum_{k=1}^{d}\binom{d}{k}\frac{5^{d-k}}{6^d - 5^d}\mathbb{E}_{d-k}[N] \\ &= \frac{6^d}{6^d - 5^d} + \sum_{j=0}^{d-1}\binom{d}{j}\frac{5^j}{6^d - 5^d}\mathbb{E}_j[N] \\ &= \frac{1}{6^d - 5^d}\left(6^d + \sum_{j=0}^{d-1}\binom{d}{j}5^j\mathbb{E}_j[N]\right). \end{aligned}$$

The result follows because $\mathbb{E}_0[N] = 0$.

b) The following Matlab function can be used to evaluate the mean of N. (It requires evaluation of the binomial coefficient as given in Listing 1.1).

```
function e = dsixes(d)
e=zeros(d,1);  % reserve memory for faster execution.
e(1)=6;
for m=2:d
  s=0;
  for k=1:m, s=s + 5^k * c(m,k) * e(k); end
  e(m)=(6^m + s) / (6^m - 5^m);
end
```

c) To do this exercise, familiarity with (among other things) the very basics of regression is required. The following handful of commands should help get the less initiated on the right track. (To discover which Matlab function performs regression, type `help stats`; to learn how it works, type `help regress`.)

```
d=50;            % that's enough for now
e = nsixes(d);  % compute expected value for n=1 through n=50
plot(1:d,e)     % have a look: it is nonlinear!
% What functions have some resemblance? log(n), 1/n and sqrt(n)
% So try one or more of them as regressors (with a constant).
% Here's an example for building a regressor matrix:
kon = ones(d,1); v=(1:d)'; X = [kon v 1./v];
[B,BINT,R,RINT,STATS] = REGRESS(e,X); B, STATS(1)
```

Some trial and error reveals that the regressors $1/d$ and $\sqrt{d}$ (and a constant) work very well, yielding (for the expectation values for $d = 150$ through $d = 200$) a parsimonious model with an R^2 measure of fit of 0.9999998 and coefficients

$$\mathbb{E}_d[N] \approx 26.4604689 + 0.554971041\sqrt{d} - 313.709051d^{-1}.$$

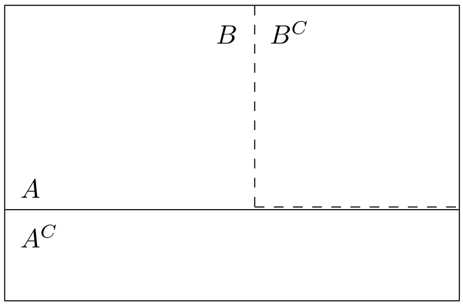
For $d = 394$ (which is well outside of the range used for "estimation"), $\mathbb{E}_{394}[N] = 0.36452$ exactly, with the approximation yielding 36.68. After $d = 394$, the previous program fails to work. For $d = 1,000$, the approximation yields 43.70; for $n = 10,000$, it gives 81.93.

d) The following Matlab function can be used to simulate the r.v. N `S` times, based on `d` dice.

```
function vec =dsixessim(d,S)
vec=zeros(S,1);
m=d;
for i=1:S
  if mod(i,100)==0, i, end % just show the progress
  z=0;
  while d>0
    w=unidrnd(6,d,1);
    x =sum(w==6);
    d=d-x;
    z=z+1;
  end
  d=m;
  vec(i)=z;
end
```

e) It is easy to see that $A^c \cap B^c = \emptyset$ so that, from De Morgan's law, $\Pr(A \cup B) = 1$. Also, A^c implies B, i.e., $A^c \subset B$, and B^c implies A, i.e., $B^c \subset A$.

The Venn diagram looks something like that in the figure below.



For $\Pr(A)$, using the hint, the probability that a single die does **not** show a **6** after x throws is clearly $(5/6)^x$. The complement (i.e., after x throws, it shows a **6**) thus has probability $1 - (5/6)^x$. Now, the independence of the dice imply

$$\Pr(A) = (1 - (5/6)^x)^d .$$

(It might help to imagine that, at each toss, *all* the dice are thrown, but you keep track of which have already displayed a **6** at least once.) Similarly, $\Pr(B) = 1 - \left(1 - (5/6)^{x-1}\right)^d$. Thus, with $p = 5/6$,

$$\Pr(N = x\,; d) = (1 - p^x)^d - \left(1 - p^{x-1}\right)^d \mathbb{I}_{\{1,2,\ldots\}}(x), \qquad d \in \mathbb{N}. \tag{S-4.1}$$

For $x = 1$, this reduces to $(1/6)^d$, which is clearly correct, while for $d = 1$,

$$\Pr(N = x\,; 1) = \left(\frac{5}{6}\right)^{x-1} - \left(\frac{5}{6}\right)^x = \frac{1}{6}\left(\frac{5}{6}\right)^{x-1} \mathbb{I}_{\{1,2,\ldots\}}(x),$$

which is just the geometric distribution and also correct. Figure S-4.1 compares the derived mass function to kernel density estimates of simulated values,[24] confirming its correctness.

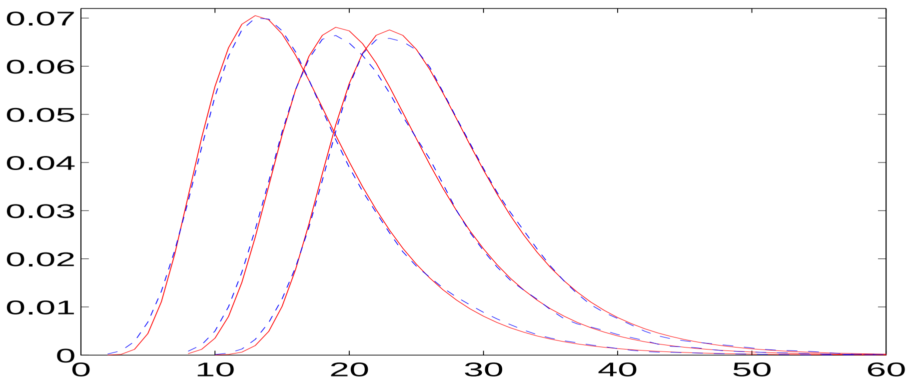


Figure S-4.1: Mass functions (S-4.1) (solid lines) and kernel density estimates of the simulated density (dashed lines) based on 10,000 replications, for $d = 10$, $d = 30$ and $d = 60$, from left to right.

[24] Kernel density estimation can be thought of as "connecting the dots" formed by the bars in a histogram, and then scaling it such that it is a proper density (i.e., that it integrates to one). More detail on this, and a program to compute it, will be given in §7.4.2.

f) Using the binomial theorem,

$$\begin{aligned}
\mathbb{E}_d[N] &= \sum_{x=1}^{\infty} x\left((1-p^x)^d - \left(1-p^{x-1}\right)^d\right) \\
&= \sum_{x=1}^{\infty} x\left(\sum_{j=0}^{d}\binom{d}{j}\left((-p^x)^j\right) - \sum_{j=0}^{d}\binom{d}{j}\left(-p^{x-1}\right)^j\right) \\
&= \sum_{x=1}^{\infty} x\left(\sum_{j=0}^{d}\binom{d}{j}(-1)^j\left(p^{jx}\right) - \sum_{j=0}^{d}\binom{d}{j}(-1)^j p^{j(x-1)}\right) \\
&= \sum_{x=1}^{\infty} x\left(\sum_{j=0}^{d}\binom{d}{j}(-1)^j p^{jx}\left(1-p^{-j}\right)\right)
\end{aligned}$$

Noting that the inner product is zero for $j=0$, switching the sums, and using the fact that, for $q=p^j$,

$$\sum_{x=1}^{\infty} x q^x = \frac{q}{(1-q)^2},$$

we have

$$\begin{aligned}
\mathbb{E}_d[N] &= \sum_{x=1}^{\infty} x\left(0+\sum_{j=1}^{d}\binom{d}{j}(-1)^j p^{jx}\left(1-p^{-j}\right)\right) \\
&= \sum_{j=1}^{d}\binom{d}{j}(-1)^j\left(1-p^{-j}\right)\sum_{x=1}^{\infty} x p^{jx} \\
&= \sum_{j=1}^{d}\binom{d}{j}(-1)^j\left(1-p^{-j}\right)\frac{p^j}{\left(1-p^j\right)^2} \\
&= \sum_{j=1}^{d}\binom{d}{j}\frac{(-1)^{j+1}}{1-p^j}.
\end{aligned}$$

Solution to Problem 4.7:

a) The number of trials performed until (and including) a 5 or 7 appears is geometric with $p = \Pr(\{X=5\}\bigcup\{X=7\}) = \Pr(X=5)+\Pr(X=7) = 10/36$ and, from Problem 4.3, expected value $p^{-1}=3.6$.

b) **Solution 1** Define E_i to be the event that neither $X=5$ or $X=7$ on the first $i-1$ trials, but $X=5$ on the i^{th} trial. As $\Pr(E_i) = \left(1-\frac{10}{36}\right)^{i-1}\frac{4}{36}$,

$$\Pr(E) = \Pr\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty}\Pr(E_i) = \frac{1}{9}\sum_{i=1}^{\infty}\left(\frac{13}{18}\right)^{i-1} = \frac{2}{5}.$$

Solution 2 Consider the first trial. Let F be the event that it results in $X=5$, G the event that $X=7$ and H the event that $X\notin\{5,7\}$, i.e., the first trial results in something other than 5 or 7. Note that events F, G and H partition the first trial without overlap, i.e., events F, G and H are exclusive and exhaustive, and that $\Pr(E\mid F)=1$, $\Pr(E\mid G)=0$ and, as trials are independent, $\Pr(E\mid H)=\Pr(E)$. It follows that

$$\Pr(E) = \Pr(F) + \Pr(E)\left(1-\Pr(F)-\Pr(G)\right),$$

which, upon solving, gives (4.13). Plugging in the values for this problem yields $\Pr(E)=2/5$.

Solution to Problem 4.8:

a) Simple calculation shows $\Pr(S = 7) = 6/36$ and $\Pr(S \leq 4) = 1/36 + 2/36 + 3/36 = 6/36$. These two events are disjoint, so that $\Pr(S = 7 \cup S \leq 4) = 1/3$ and, from the nature of the trials and stopping procedure, $N \sim \text{Geo}(1/3)$ with density (4.18) and, from Problem 4.3, $\mathbb{E}[N] = 3$ and $\mathbb{V}(N) = 6$.

b) Because the two events are equally likely, from (4.13), the answer is 1/2.

Solution to Problem 4.9: From (4.10),

$$\Pr(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}\frac{1}{1-e^{-\lambda}} = \frac{\lambda^x}{(e^{\lambda}-1)\,x!}$$

and

$$\mathbb{E}[X] = \frac{1}{e^{\lambda}-1}\sum_{x=1}^{\infty} x\frac{\lambda^x}{x!} = \frac{\lambda}{e^{\lambda}-1}\sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!} = \frac{\lambda}{e^{\lambda}-1}e^{\lambda} = \frac{\lambda}{1-e^{-\lambda}}$$

and, as

$$\mathbb{E}[X(X-1)] = \frac{1}{e^{\lambda}-1}\sum_{x=1}^{\infty} x(x-1)\frac{\lambda^x}{x!} = \frac{\lambda^2}{e^{\lambda}-1}\sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{(x-2)!} = \frac{\lambda^2}{1-e^{-\lambda}},$$

we have

$$\mathbb{V}(X) = \mathbb{E}[X(X-1)] + \mathbb{E}[X] - (\mathbb{E}[X])^2 = \frac{\lambda\left(1-\lambda e^{-\lambda}-e^{-\lambda}\right)}{\left(1-e^{-\lambda}\right)^2}.$$

Solution to Problem 4.10: This is

$$1\cdot\frac{3}{5} + 2\cdot\frac{2}{5}\frac{3}{4} + 3\cdot\frac{2}{5}\frac{1}{4} = 1.5.$$

Solution to Problem 4.11:

a) Denote a defective as D and a nondefective as G. Notice that, for $T = 3$, the event $\{GGG\}$ also identifies the 2 defectives as being the 2 components remaining. Thus,

$$\begin{aligned}
\Pr(T=2) &= \Pr(DD) = \frac{2}{5}\frac{1}{4} = \frac{1}{10}\\
\Pr(T=3) &= \Pr(GDD) + \Pr(DGD) + \Pr(GGG) = \frac{3}{5}\frac{2}{4}\frac{1}{3} + \frac{2}{5}\frac{3}{4}\frac{1}{3} + \frac{3}{5}\frac{2}{4}\frac{1}{3}\\
&= \frac{3}{10}\\
\Pr(T=4) &= 1 - \Pr(T=2) - \Pr(T=3) = \frac{6}{10},
\end{aligned}$$

and zero otherwise. Also,

$$\mathbb{E}[T] = 2\cdot\frac{1}{10} + 3\cdot\frac{3}{10} + 4\cdot\frac{6}{10} = 3.5.$$

b) This is similar to the $N = 5$ case, again noticing that, for $T = N - 2$, there is the additional possibility that all $N - 2$ chosen items are nondefective, so that one knows that the remaining

2 are the defective ones. We have

$$\begin{aligned}
\Pr(T=2) &= \Pr(DD) = \frac{2}{N}\frac{1}{N-1} = \binom{N}{2}^{-1} \\
\Pr(T=3) &= \Pr(GDD) + \Pr(DGD) \\
&= \frac{N-2}{N}\frac{2}{N-1}\frac{1}{N-2} + \frac{2}{N}\frac{N-2}{N-1}\frac{1}{N-2} = 2\binom{N}{2}^{-1} \\
\Pr(T=4) &= \Pr(GGDD) + \Pr(GDGD) + \Pr(DGGD) = 3\binom{N}{2}^{-1} \\
&\vdots \\
\Pr(T=N-3) &= (N-4)\binom{N}{2}^{-1} \\
\Pr(T=N-2) &= (N-3)\binom{N}{2}^{-1} + \Pr\Bigg(\underbrace{GG\cdots G}_{N-2\text{ times}}\Bigg) \\
&= (N-3)\binom{N}{2}^{-1} + \frac{\binom{N-2}{N-2}\binom{2}{0}}{\binom{N}{2}} = (N-2)\binom{N}{2}^{-1}
\end{aligned}$$

and

$$\begin{aligned}
\Pr(T=N-1) &= 1 - \sum_{i=2}^{N-2}\Pr(T=i) \\
&= 1 - \binom{N}{2}^{-1}(1+2+\cdots+(N-4)+(N-2)) \\
&= 1 - \binom{N}{2}^{-1}\left(1+\frac{(N-3)(N-2)}{2}\right) \\
&= 1 - \frac{2+(N-3)(N-2)}{N(N-1)} = 4\frac{N-2}{N(N-1)} \\
&= 2(N-2)\binom{N}{2}^{-1}
\end{aligned}$$

so that

$$f_T(t;N) = \binom{N}{2}^{-1}\begin{cases} (t-1) & \text{if } 2 \le t \le N-3 \\ N-2 & \text{if } t = N-2 \\ 2(N-2) & \text{if } t = N-1 \\ 0 & \text{otherwise.} \end{cases}$$

c) Using the rules of summation discussed in Example 1.18,

$$\begin{aligned}
\mathbb{E}[T] &= \sum_{t=2}^{N-3} t f_T(t) = \binom{N}{2}^{-1}\left(\sum_{t=2}^{N-3} t(t-1) + (N-2)^2 + 2(N-2)(N-1)\right) \\
&= \frac{2(N-2)(N+2)}{3(N-1)}
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}\left[T^2\right] &= \binom{N}{2}^{-1}\left(\sum_{t=2}^{N-3} t^2(t-1) + (N-2)^3 + 2(N-2)(N-1)^2\right) \\
&= \frac{3N^4+2N^3-27N^2+10N+24}{6N(N-1)}
\end{aligned}$$

so that

$$\begin{aligned}\mathbb{V}(T) &= \mathbb{E}\left[T^2\right] - \left(\mathbb{E}[T]\right)^2 \\ &= \frac{3N^4 + 2N^3 - 27N^2 + 10N + 24}{6N(N-1)} - \left(\frac{2(N-2)(N+2)}{3(N-1)}\right)^2 \\ &= \frac{N^5 - 3N^4 - 23N^3 + 111N^2 - 86N - 72}{18N(N-1)^2}.\end{aligned}$$

d) Clearly, $\mathbb{E}[t] \to \frac{2}{3}N$, which is reasonable because, on average, the 2 defectives will partition the random "row" of components into 3 sections, so that one must sample 2/3 of them to reach the 2$^{\text{nd}}$ defective. See §6.4.3 for a more formal presentation of this argument.

e) Listing S-4.1 shows one way.

f) In order for $T = N - 1$, it must be the case that a defective and a nondefective are the last two items remaining whose status is unknown and one of which will be inspected. There are $\binom{N-2}{D-1}\binom{2}{1}$ out of $\binom{N}{D}$ ways which this can occur.

From the plots, it appears as though the probability decays geometrically. With $X = N - T$, we analytically know $f_X(1) = \Pr(X = 1)$. For $f_X(x)$, $x \geq 2$, we could take $f_X(x) \approx p(1-p)^{x-2}$. For example, with $N = 48$ and $D = N/4 = 12$, simulation gives the empirical values 0.30, 0.20, 0.14, 0.10, 0.06 and 0.04 for $X = 2$ through $X = 7$. Taking $p = 0.30$ gives approximate probabilities of 0.300, 0.210, 0.147, 0.103, 0.072 and 0.050, which are indeed quite close to the empirical values. (The *estimate* of p, denoted $\hat{p}$, can be obtained as $1/\bar{X}$, which gave 0.3005 in this case. Estimation in general, and a discussion of this estimator for p, will be discussed in a later chapter.)

One might guess that the geometric approximation works well for large values of N and values of D which grow with N, e.g., $D = N/4$. The ambitious reader could try to verify this by simulation for a large variety of N and D and attempt to further approximate $\hat{p}$ as a simple function of N and D.

Solution to Problem 4.12:

a) Taking $X_i = N_i - K$, $x_i = N_i - k$, $i = 1, 2$ and $N = N_1 + N_2$ shows that

$$\begin{aligned} f(k; N_1, N_2, p) &= \Pr(K = k \mid N_1, N_2, p) \\ &= \Pr(X_2 = x \cap \text{lhpe}) + \Pr(X_1 = x \cap \text{rhpe}) \\ &= \binom{N-k}{N_1} p^{N_1+1} (1-p)^{N_2-k}\, \mathbb{I}_{\{0,1,\ldots,N_2\}}(k) \\ &\quad + \binom{N-k}{N_2} (1-p)^{N_2+1}\, p^{N_1-k} \mathbb{I}_{\{0,1,\ldots,N_1\}}(k). \end{aligned} \qquad \text{(S-4.2)}$$

A Matlab program which accomplishes the desired task is shown in Listing S-4.2. It is "restricted" to having $N_1 \geq N_2$.

b) Listing S-4.3 shows one possible way of programming this.

Solution to Problem 4.13: We have that $\mathbb{E}_N[K]$ is given by

$$\begin{aligned} \sum_{j=0}^{N} j \binom{2N-j}{N}\left(\frac{1}{2}\right)^{2N-j} &= \sum_{k=0}^{N-1} (k+1) \binom{2N-k-1}{N}\left(\frac{1}{2}\right)^{2N-k-1} \\ = \quad \sum_{k=0}^{N-1} (k+1)\binom{2N-k-2}{N-1}\left(\frac{1}{2}\right)^{2N-k-1} &+ \sum_{k=0}^{N-1} (k+1)\binom{2N-k-2}{N}\left(\frac{1}{2}\right)^{2N-k-1} \\ =: \quad G + H. \end{aligned}$$

```
function [e,v,tests]=electriciden(n,d,sim);
e=0; esq=0; if nargout>=3, tests=zeros(sim,1); end
for i=1:sim;
    a=electricsim(n,d,0);
    e=e+a; esq=esq+a.^2;
    if nargout>=3, tests(i)=a; end
end;
e=e/sim; esq=esq/sim; v=esq-e.*e;
if d==2 % for this we have an exact solution
  true_e = 2*(n-2)*(n+2)/3/(n-1)
  true_v = (n^5 - 3*n^4 - 23*n^3 + 111*n^2 -86*n -72) / (18*n*(n-1)^2)
end

function a=electricsim(n,d,zaehler);
a=0;
k=n-d;     % k is number of nondefectives
if k>=1;   % If at least one nondefective remaining, keep drawing
  if d>=1; % If at least one defective remaining, keep drawing
    y=unifrnd(0,1);
    a=zaehler;
    zaehler=zaehler+1;
    if y<=(d/n); a=electricsim(n-1,d-1,zaehler); % draw a defective
    else, a=electricsim(n-1,d,zaehler);  % draw a nondefective
    end;
  else
    if zaehler>a; a=zaehler; end;
  end;
else
  if zaehler>a; a=zaehler; end;
end;
```

Program Listing S-4.1: Parameters `n` and `d` correspond to N and D, while `sim` is the desired number of replications to perform. Observe how the subprogram `electricsim` uses recursion (it calls itself) and keeps track of the number of required inspections via variable zaehler (Zähler being German for counter).

```
function vec=banach(n1,n2,p,sim)
vec=zeros(sim,1);
for i=1:sim
  vec(i)=simul(n1,n2,p); if mod(i,100)==0, i, end
end
tt=tabulate(vec+1); a=tt(:,2); mx=tt(end,1)-1; b=0:mx; a=a./sim;
true=echt(n1,n2,p); plot(b,a,'r-',0:mx,true(1:mx+1),'go')
mn=min(vec); ax=axis; axis([mn mx 0 ax(4)]), set(gca,'fontsize',14)

function echte=echt(n1,n2,p) % true mass function
mx=max(n1,n2); echte=zeros(1,mx); n=n1+n2; d=0:mx;
d1=0:n2; h1=c(n-d1,n1);
k1=p^(n1+1).*(1-p).^(n2-d1); h1=h1.*k1;
d2=0:n1; h2=c(n-d,n2);
k2=(1-p)^(n2+1).*p.^(n1-d2); h2=h2.*k2;
if n1>n2
  g=zeros(1,n1-n2); h1=[h1,g];
else
  g=zeros(1,n2-n1); h2=[h2,g];
end
echte=h1+h2;

function output=simul(n1,n2,p);
ok=1; zaehlerone=0; zaehlertwo=0;
while ok
  y=unifrnd(0,1,1,1);
  if y<p  % box 1
    if n1==0, x=n2; ok=0; end
    if n1>0, n1=n1-1; zaehlerone=zaehlerone+1; end
  end
  if y>p
    if n2==0, x=n1; ok=0; end
    if n2>0, n2=n2-1; zaehlertwo=zaehlertwo+1; end
  end
end
output=x;
```

Program Listing S-4.2: Plots (S-4.2) overlaid with simulated values

```
function minmax=banachmultisim(N,p,sim)
if any(p<=0) | any(p>=1) | sum(p)~=1, error('bad p'), end
w=max(N); y=zeros(2,sim);
for i=1:sim
  [s,d]=banachmulti(N,p);
  y(1:2,i)=[s,d]';
end
tt1=tabulate(y(1,1:sim)+1); tt2=tabulate(y(2,1:sim)+1);
a=tt1(:,2); b=tt2(:,2); mx=tt1(end,1)-1; mi=tt2(end,1)-1;
a=a./sim; b=b./sim;
if length(N)==2
  plot(0:mx,a,'r-'), title('Mass Function of Remaining Matches')
else
  plot(0:mx,a,'r-',0:mi,b,'b--')
  title('Marginal Mass Function of Minimum and Maximum Remaining Matches')
end
minmax=y';

function [ma,mi]=banachmulti(x,p);
n=length(x); ok=1; mi=min(x);
while ok==1
  s=0; zaehler=0; r=unifrnd(0,1,1,1);
  for i=1:n, if zaehler==0
      s=s+p(i);
      if r<s, zaehler=i; end
  end, end
  x(zaehler)=x(zaehler)-1; mi=min(x); ma=max(x);
  if mi<1, ok=0; end
end
help=find(x==0); x(help)=max(x)+1; mi=min(x);
```

Program Listing S-4.3: Plots the mass function of the min and max of the remaining matches in the "vector" Banach matchbox problem

Thus, G is given by

$$\begin{aligned} & \frac{1}{2}\left\{\sum_{k=0}^{N-1} k\binom{2(N-1)-k}{N-1}\left(\frac{1}{2}\right)^{2(N-1)-k}+\sum_{k=0}^{N-1}\binom{2(N-1)-k}{N-1}\left(\frac{1}{2}\right)^{2(N-1)-k}\right\} \\ = \; & \frac{1}{2}\mathbb{E}_{N-1}[K]+\frac{1}{2} \end{aligned}$$

and, with $j=k+2$, H is

$$\begin{aligned} & \sum_{j=2}^{N}(j-1)\binom{2N-j}{N}\left(\frac{1}{2}\right)^{2N-j+1} \\ = \; & \frac{1}{2}\left\{\sum_{j=2}^{N} j\binom{2N-j}{N}\left(\frac{1}{2}\right)^{2N-j}-\sum_{j=2}^{N}\binom{2N-j}{N}\left(\frac{1}{2}\right)^{2N-j}\right\} \\ = \; & \frac{1}{2}\left\{\mathbb{E}_N[K]-\binom{2N-1}{N}\left(\frac{1}{2}\right)^{2N-1}\right\}-\frac{1}{2}\left\{1-\binom{2N}{N}\left(\frac{1}{2}\right)^{2N}-\binom{2N-1}{N}\left(\frac{1}{2}\right)^{2N-1}\right\} \\ = \; & \frac{1}{2}\mathbb{E}_N[K]-\frac{1}{2}+\binom{2N}{N}\left(\frac{1}{2}\right)^{2N+1}. \end{aligned}$$

Simplifying $G+H$ yields the recursion

$$\mathbb{E}_N[K]=\mathbb{E}_{N-1}[K]+\binom{2N}{N}\left(\frac{1}{2}\right)^{2N}, \quad \mathbb{E}_1[K]=\frac{1}{2}.$$

Solving and using (1.10), (1.9) and (1.6) to simplify yields

$$\mathbb{E}_N[K]=\sum_{i=1}^{N}\binom{2i}{i}\left(\frac{1}{2}\right)^{2i}=\sum_{i=1}^{N}(-1)^i\binom{-\frac{1}{2}}{i}=\sum_{i=1}^{N}\binom{i-1/2}{i}=\binom{N+1/2}{N}-1.$$

Finally, using the approximation to $\binom{N+1/2}{N}$ in (A.95) yields

$$\mathbb{E}_N[K]\approx -1+\frac{2N+1}{\sqrt{N\pi}}\approx -1+2\sqrt{N/\pi}.$$

Solution to Problem 4.14: If this holds for n, then, for $n+1$,

$$\begin{aligned} \mathbb{E}[M \mid n+1] &= \sum_{m=1}^{n} m\, c_{m,n+1}+\sum_{m=n+1}^{n+1} m\, c_{m,n+1} \\ &= \sum_{m=1}^{n}\frac{1}{(m-1)!}\sum_{i=0}^{n+1-m}\frac{(-1)^i}{i!}+\frac{1}{n!} \\ &= \mathbb{E}[M \mid n]+\sum_{m=1}^{n}\frac{1}{(m-1)!}\frac{(-1)^{n+1-m}}{(n+1-m)!}+\frac{1}{n!}, \end{aligned}$$

so that we must show that

$$\sum_{m=1}^{n}\frac{1}{(m-1)!}\frac{(-1)^{n+1-m}}{(n+1-m)!}=-\frac{1}{n!}, \quad \text{or} \quad 1=\sum_{m=1}^{n}\frac{n!}{(m-1)!}\frac{(-1)^{n-m}}{(n+1-m)!}.$$

But, with $j=m-1$ and repeated use of (1.4),

$$\sum_{m=1}^{n}\frac{n!}{(m-1)!}\frac{(-1)^{n-m}}{(n+1-m)!}$$

is equal to

$$
\begin{aligned}
& \sum_{m=1}^{n} \binom{n}{m-1} (-1)^{n-m} = \sum_{j=0}^{n-1} \binom{n}{j} (-1)^{n-j-1} \\
= \; & \sum_{j=0}^{n-1} \binom{n-1}{j} (-1)^{n-j-1} + \sum_{j=1}^{n-1} \binom{n-1}{j-1} (-1)^{n-j-1} \\
= \; & (1-1)^{n-1} + \sum_{j=1}^{n-1} \binom{n-2}{j-1} (-1)^{n-j-1} + \sum_{j=2}^{n-1} \binom{n-2}{j-2} (-1)^{n-j-1} \\
= \; & (1-1)^{n-1} + (1-1)^{n-1} + \cdots \\
= \; & \vdots \\
= \; & \sum_{j=n-1}^{n-1} \binom{n-(n-1)}{j-(n-1)} (-1)^{n-j-1} \\
= \; & (-1)^{n-(n-1)-1} = (-1)^0 = 1.
\end{aligned}
$$

Solution to Problem 4.15: With $p = 1 - \lambda/r$,

$$
\begin{aligned}
f_X(x; r, p) &= \binom{r+x-1}{x} p^r (1-p)^x \, \mathbb{I}_{\{0,1,\ldots\}}(x) \\
&= \underbrace{\frac{(r+x-1)!}{(r-1)!\; r^x}}_{\to 1} \underbrace{\left(1 - \frac{\lambda}{r}\right)^r}_{\to e^{-\lambda}} \frac{\lambda^x}{x!} \mathbb{I}_{\{0,1,\ldots\}}(x)
\end{aligned}
$$

so that $X \overset{\text{asy}}{\sim} \text{Poi}(\lambda)$.

Solutions to Chapter 5: Multivariate Random Variables

Solution to Problem 5.1:

$$\begin{aligned}\Pr\left(X_1 > b_1,\, X_2 > b_2\right) &= \Pr\left(b_1 < X_1 < \infty,\, b_2 < X_2 < \infty\right)\\ &= 1 + F_{X_1,X_2}\left(b_1, b_2\right) - F_{X_1}\left(b_1\right) - F_{X_2}\left(b_2\right).\end{aligned}$$

Solution to Problem 5.2: From the nonnegativity of the variance and (6.4),

$$0 \leq \mathbb{V}\left(\frac{X_1}{\sigma_1} + \frac{X_2}{\sigma_2}\right) = \frac{\mathbb{V}\left(X_1\right)}{\sigma_1^2} + \frac{\mathbb{V}\left(X_2\right)}{\sigma_2^2} + 2\frac{\mathrm{Cov}\left(X_1 X_2\right)}{\sigma_1\sigma_2} = 2\left(1+\rho\right),$$

where $\rho = \mathrm{Corr}\left(X_1, X_2\right)$, so that $-1 \leq \rho$. Similarly,

$$0 \leq \mathbb{V}\left(\frac{X_1}{\sigma_1} - \frac{X_2}{\sigma_2}\right) = \frac{\mathbb{V}\left(X_1\right)}{\sigma_1^2} + \frac{\mathbb{V}\left(X_2\right)}{\sigma_2^2} - 2\frac{\mathrm{Cov}\left(X_1 X_2\right)}{\sigma_1\sigma_2} = 2\left(1-\rho\right),$$

implying that $\rho \leq 1$. Together, this yields that $\left|\mathrm{Corr}\left(X_1, X_2\right)\right| \leq 1$.

Solution to Problem 5.3: For the inner integral in

$$I \equiv 2\int\int_{x<y} F(x)(1-F(y))dxdy,$$

take $u = F\left(x\right)$ and $dv = dx$, so that

$$\begin{aligned}\int_{-\infty}^{y} F(x)dx &= F(x)x|_{-\infty}^{y} - \int_{-\infty}^{y} xf(x)dx\\ &= yF(y) - \int_{-\infty}^{y} xf(x)dx = \int_{-\infty}^{y}(y-x)f(x)dx\end{aligned}$$

and substituting,

$$I = 2\int\int_{x<y} f(x)(y-x)(1-F(y))dxdy.$$

Integrating this expression with respect to y, using $u = 1 - F\left(y\right)$ and $dv = \left(y - x\right)$ gives

$$\begin{aligned}\int_{x}^{\infty}(y-x)(1-F(y))dy &= \left\{\frac{1}{2}(y-x)^2(1-F(y))\right\}\Big|_{x}^{\infty} + \frac{1}{2}\int_{x}^{\infty}(y-x)^2 f(y)dy\\ &= \frac{1}{2}\int_{x}^{\infty}(y-x)^2 f(y)dy.\end{aligned}$$

Thus,

$$\begin{aligned}I &= 2\times\frac{1}{2}\int\int_{x<y}(y-x)^2 f(x)f(y)dxdy\\ &= \frac{1}{2}\int\int(y-x)^2 f(x)f(y)dxdy = \frac{1}{2}\mathbb{E}\left[\left(X_1 - X_2\right)^2\right].\end{aligned}$$

Solution to Problem 5.4: From the definitions,

$$\mathbb{E}[P] = \Pr(P = 1) = \Pr(\text{at least one } B_i \text{ is one}) = \Pr\left(\bigcup_{i=1}^{n} A_i\right),$$

so that

$$\begin{aligned}
\Pr\left(\bigcup_{i=1}^{n} A_i\right) &= \mathbb{E}[P] = \mathbb{E}\left[1 - \prod_{i=1}^{n}(1 - B_i)\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{n} B_i - \sum_{i<j} B_i B_j - \cdots + (-1)^{n+1} B_1 \cdots B_n\right] \\
&= \sum_{i=1}^{n} \Pr(A_i) - \sum_{i<j} \Pr(A_i A_j) - \cdots + (-1)^{n+1} \Pr(A_1 \cdots A_n).
\end{aligned}$$

Solution to Problem 5.5: Starting with the boundary conditions, we have $R_{0d} = 1$ for $d > 0$ and $R_{s0} = 0$ for $s > 0$. Then, for s and d both positive,

$$R_{sd} = pR_{s-1,d} + qR_{s,d-1}.$$

Solution to Problem 5.6: Program `egg` accomplishes this, given in Listings S-5.1 and S-5.2, with sample output shown in Figures S-5.1, S-5.2 and S-5.3. Note the use of Matlab's `find` and `tabulate` functions in function `getfreqs`, which is necessary because not all $n \in [r, \max(\texttt{needed})]$ will be observed in the simulated values.

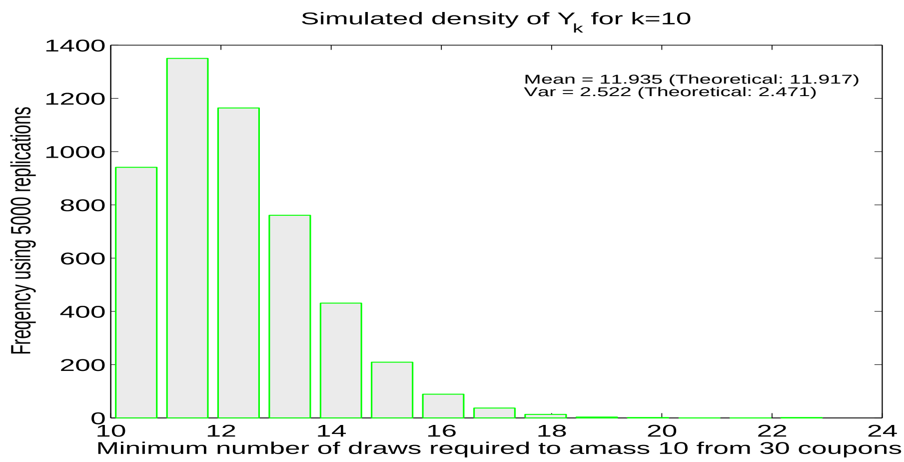


Figure S-5.1: Histogram output from `egg(30,10)`

```
function egg(r,k,reps,seed)
if nargin<2, k=r; end
if nargin<3, reps=5000; end
if nargin<4, seed=9999; end
rand('seed',seed); needed=zeros(reps,1);

for i=1:reps
  toy=zeros(r,1);  % boolean vector of r toys
  count=0;         % how many toys purchased.
  while sum(toy)<k
    newtoy = unidrnd(r,1,1); % same as ceil(N * rand(1,1));
    toy(newtoy)=1; count=count+1;
  end
  needed(i)=count;
end

hist(needed,range(needed)+1); set(gca,'fontsize',16)
title (['Simulated density of Y_k for k=',int2str(k)],'fontsize',16);
ylabel (['Freqency using ', int2str(reps),' replications'],'fontsize',16)
xlabel (['Minimum number of draws required to amass ',int2str(k), ...
        ' from ',int2str(r),' coupons'],'fontsize',16);

i=0:k-1; mn= r*sum(1./(r-i)); vr= r*sum(i./(r-i).^2);
outm = ['Mean = ',sprintf('%0.3f',mean(needed)), ...
       ' (Theoretical: ',sprintf('%0.3f',mn),')'];
outv = ['Var = ',sprintf('%0.3f',var(needed)), ...
       ' (Theoretical: ',sprintf('%0.3f',vr),')'];
ax=axis;
text( 1*ax(1)/2+ax(2)/2, 10*ax(4)/11, outm, 'fontsize', 12)
text( 1*ax(1)/2+ax(2)/2, 7*ax(4)/8, outv, 'fontsize', 12)

if k==r  % calculate true pdf values.
  uplim=max(needed)+round(sqrt(r/2)); % extend to compare to true probs.
  cdf=[]; for n=(r-1):uplim, cdf=[cdf F(n,r)]; end % see function F below
  pdf=diff(cdf); % f(X=n) = F(X=n) - F(X=n-1) and F(X=N-1)=0.
  emppmf = getfreqs(r,uplim,needed)/reps; % now it sums to one.
  figure(2), plot(r:uplim,emppmf-pdf), grid, set(gca,'fontsize',16)
end
```

Program Listing S-5.1: Simulate pmf of Y_k and measure discrepancy between simulated and actual pmf. The program is continued below in Listing S-5.2, where function `getfreqs` is given.

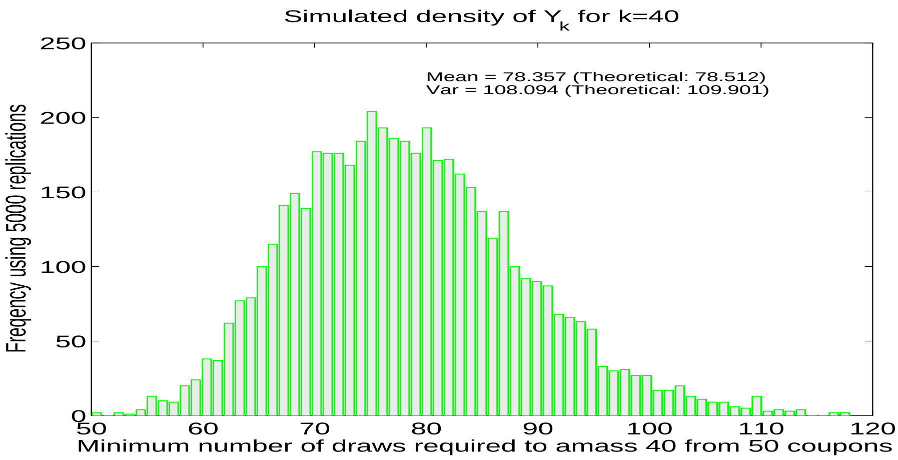


Figure S-5.2: Histogram output from `egg(50,40)`

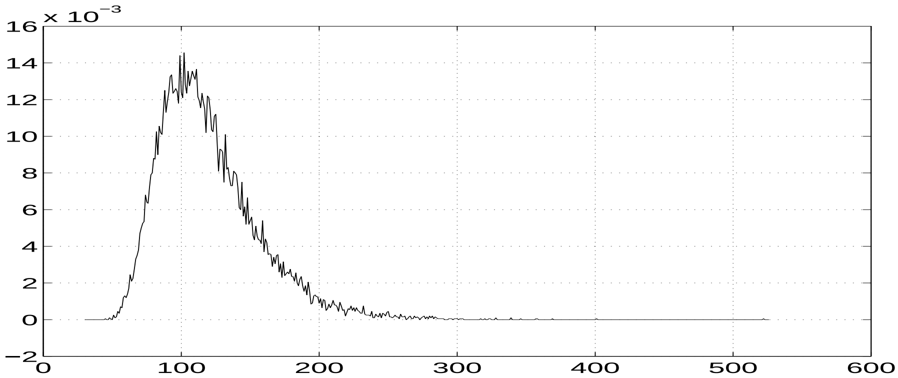


Figure S-5.3: Density comparison output from `egg(30,30,20000)`

```
function out=getfreqs(r,uplim,needed);
% necessary because some values of n may not occur
out=zeros(1,uplim-r+1);
tab=tabulate(needed);      % see below
for i=r:uplim
  pos=find(tab(:,1)==i);   % did we observe n=i?
  if length(pos)>0         % yes, observed at least one value,
    out(i-r+1)=tab(pos,2); %   so record it in out
  end                      % otherwise, it stays zero.
end

function P=F(n,r)
  i=0:(r-1);
  P=sum( gammaln(r+1) - gammaln(r-i+1) ...
      - gammaln(i+1) + n*log((r-i)/r) + i*log(-1) );
  P=real(exp(P));
```

Program Listing S-5.2: Listing S-5.1 continued. `tabulate` is a built–in Matlab function with output as two columns, the first is the value, the second is the frequency

Solutions to Chapter 6: Sums of Random Variables

Solution to Problem 6.1:

a) Letting $n = n_1 + n_2$,

$$\begin{aligned}
\Pr(Y = y) &= \sum_{i=-\infty}^{\infty} \Pr(X_1 = i)\Pr(X_2 = y - i) \\
&= \sum_{i=0}^{y} \binom{n_1}{i} p^i (1-p)^{n_1 - i} \binom{n_2}{y-i} p^{y-i} (1-p)^{n_2 - y + i} \\
&= p^y (1-p)^{n_1 + n_2 - y} \sum_{i=0}^{y} \binom{n_1}{i}\binom{n_2}{y-i} \\
&= \binom{n}{y} p^y (1-p)^{n-y},
\end{aligned}$$

using the combinatoric identity (1.28).

b) With $r = r_1 + r_2$,

$$\begin{aligned}
\Pr(Y = y) &= \sum_{i=-\infty}^{\infty} \Pr(X_1 = i)\Pr(X_2 = y - i) \\
&= \sum_{i=0}^{y} \binom{r_1 + i - 1}{i} p^{r_1} (1-p)^i \binom{r_2 + (y-i) - 1}{y-i} p^{r_2} (1-p)^{y-i} \\
&= p^{r_1 + r_2} (1-p)^y \sum_{i=0}^{y} \binom{r_1 + i - 1}{i}\binom{r_2 + y - i - 1}{y-i} \\
&= \binom{r + y - 1}{y} p^r (1-p)^y,
\end{aligned}$$

using the combinatoric identity (1.58). Another way of verifying this identity is to recall (6.26), i.e.,

$$\binom{w+b}{w} = \sum_{x=k}^{b+k} \binom{x-1}{k-1}\binom{w+b-x}{w-k},$$

and substitute $i = x - k$, use that $\binom{a+b}{a} = \binom{a+b}{b}$, then substitute $b = y$, $w = r_1 + r_2$ and, finally, $k = r_1$. Doing so yields

$$\binom{r_1 + r_2 + y}{y} = \sum_{i=0}^{y} \binom{i + r_1 - 1}{i}\binom{y - i + r_2}{y-i},$$

which is the desired identity with r_2 replaced by $r_2 - 1$.

Solution to Problem 6.2:

a) If N is the number of rolls needed, $N = \sum_{i=1}^{3} X_i$, where the X_i are independent geometric variables with probability $p = \frac{3}{6}, \frac{2}{6}$ and $\frac{1}{6}$, respectively, so that $\mathbb{E}[N] = 2 + 3 + 6 = 11$.

b) Now, $X_1 = 1$, while X_2 and X_3 are independent geometric variables with respective probabilities $p = \frac{5}{6}$ and $\frac{4}{6}$, so that $\mathbb{E}[N] = 1 + 6/5 + 6/4 = 3.7$.

Solution to Problem 6.3:

a) n is just $\binom{r}{2}$ or $r(r-1)/2$, and p is the probability that 2 people have the same birthday, or $1/365$.

b) (a) The events are independent because they involve different people from the group, and $\Pr(E_{1,2} \mid E_{3,4}) = \Pr(E_{1,2}) = 1/365$. (b) Because we condition on $E_{1,2}$ this is just the probability that person 3 has the same birthday as person 1, irrespective of the fact that persons 1 and 2 have the same birthday. Thus $\Pr(E_{1,3} \mid E_{1,2}) = 1/365$. (c) Clearly persons 2 and 3 have the same birthday, so $\Pr(E_{2,3} \mid E_{1,2} \cap E_{1,3}) = 1$. Thus, trials are pairwise independent, but not otherwise.

c) Using the complement,

$$p^* = 1 - \Pr(\text{no birthdays in common}) \approx 1 - \frac{e^{-\lambda}\lambda^0}{0!} = 1 - e^{-\lambda}$$

where $\lambda = np = [r(r-1)/2] / 365$. Solving yields

$$r = \left\lceil 0.5 + 0.5\sqrt{(1 - 2920\ln(1-p^*))} \right\rceil,$$

the same answer in part 12 of Example 6.2.1.

d) Now we have $\binom{r}{3} = \frac{r(r-1)(r-2)}{6}$ trials, with probability $p = 365^{-2}$, so that, as before,

$$0.5 = 1 - \Pr(\text{no birthdays in common}) \approx 1 - \frac{e^{-\lambda}\lambda^0}{0!},$$

where $\lambda = np = [r(r-1)(r-2)/6] / 365^2$ or

$$0.693 = \log(2) = \frac{r(r-1)(r-2)/6}{365^2}.$$

Instead of solving, note that $\frac{84(84-1)(84-2)/6}{365^2} = 0.715$ and $\frac{83(83-1)(83-2)/6}{365^2} = 0.690$, so that $r = 84$ is the smallest integer for which $p_{r,3} \geq 0.5$ is satisfied.

Solution to Problem 6.4:

a) Define X_i to be

$$X_i = \begin{cases} 1, & \text{if } i^{\text{th}} \text{ pair is man-woman} \\ 0, & \text{otherwise} \end{cases}$$

for $i = 1, \ldots, 6$. Then

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{6} X_i\right] = \sum_{i=1}^{6} \mathbb{E}[X_i] = \sum_{i=1}^{6} \Pr(X_i = 1) = 6\mathbb{E}[X_1],$$

because the probability is the same for each pair. To calculate $\mathbb{E}[X_1]$, consider just one man. He has a $6/11$ chance of being paired with a woman, so that $\mathbb{E}[X_1] = \frac{6}{11}$ and, thus,

$$\mathbb{E}[X] = \frac{36}{11}. \tag{S-6.1}$$

Write the variance as

$$\mathbb{V}(X) = \mathbb{V}\left(\sum_{i=1}^{6} X_i\right) = \sum_{i=1}^{6} \mathbb{V}(X_i) + \sum_{i \neq j} \operatorname{Cov}(X_i X_j),$$

and note that, for $i \neq j$,

$$\mathbb{E}[X_i X_j] = \Pr(X_i = 1, X_j = 1) = \Pr(X_j = 1 \mid X_i = 1)\Pr(X_i = 1) = \frac{5}{9}\frac{6}{11},$$

so that

$$\operatorname{Cov}(X_i X_j) = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\,\mathbb{E}[X_j] = \frac{5}{9}\frac{6}{11} - \left(\frac{6}{11}\right)^2.$$

The X_i are boolean random variables, so $\mathbb{V}(X_i) = \frac{6}{11}\left(1 - \frac{6}{11}\right)$, and

$$\mathbb{V}(X) = 6 \cdot \frac{6}{11}\left(1 - \frac{6}{11}\right) + 6 \cdot 5\left[\frac{5}{9}\frac{6}{11} - \left(\frac{6}{11}\right)^2\right] = \frac{380}{121} \approx 1.653. \qquad \text{(S-6.2)}$$

b) As in the previous question, for a given man, he has a $\frac{1}{11}$ chance of getting his wife, so $\mathbb{E}[X] = \frac{6}{11}$. Similarly,

$$\mathbb{V}(X) = 6\frac{1}{11}\left(1 - \frac{1}{11}\right) + 6 \cdot 5\left[\frac{1}{9}\frac{1}{11} - \left(\frac{1}{11}\right)^2\right] = \frac{2030}{363} \approx 5.59.$$

Solution to Problem 6.5:

$$\begin{aligned}
\mu'_{[g]} &= \sum_{r=1}^{N} f_R(r)\, r(r-1)\cdots(r-g+1) \\
&= \frac{1}{\binom{Z+N}{N}} \sum_{r=1}^{N} \binom{N-1}{r-1} \frac{(Z+1)\,Z\,(Z-1)\cdots(Z-r+2)}{r!} r(r-1)\cdots(r-g+1) \\
&= \frac{1}{\binom{Z+N}{N}} \sum_{r=1}^{N} \binom{N-1}{r-1} \frac{(Z+1)\,Z\,(Z-1)\cdots(Z-r+2)}{(r-g)!}
\end{aligned}$$

and multiplying this by

$$1 = \frac{(r-g)!}{(Z+1-g)\cdots(Z+2-r)}\binom{Z+1-g}{r-g}$$

and cancelling terms,

$$\mu'_{[g]} = \frac{(Z+1)_{[g]}}{\binom{Z+N}{N}} \sum_{r=1}^{N} \binom{N-1}{r-1}\binom{Z+1-g}{r-g},$$

where the sum could go from $r = g$ instead of $r = 1$. This can be simplified further by using (1.55), i.e.,

$$\binom{Q+M}{k+M} = \sum_{i=0}^{M} \binom{Q}{k+i}\binom{M}{i}.$$

With $i = r - 1$,

$$\begin{aligned}
\sum_{r=1}^{N} \binom{N-1}{r-1}\binom{Z+1-g}{r-g} &= \sum_{i=0}^{N-1} \binom{N-1}{i}\binom{Z+1-g}{i+1-g} \\
&= \binom{(Z+1-g)+(N-1)}{(1-g)+(N-1)} \\
&= \binom{Z+N-g}{N-g},
\end{aligned}$$

so that

$$\mu'_{[g]} = \frac{(Z+1)_{[g]}}{\binom{Z+N}{Z}}\binom{N+Z-g}{N-g},$$

as was to be shown.

Solution to Problem 6.6: From the definition of expected value and using (6.12) for the three cases $r_0 = r_1$, $r_0 = r_1 - 1$ and $r_0 = r_1 + 1$,

$$\begin{aligned}
\mu'_{[g_0,g_1]} &= \sum_{r_0=1}^{Z}\sum_{r_1=1}^{N} f_{R_0,R_1}(r_0,r_1)(r_0-1)_{[g_0]}(r_1-1)_{[g_1]} \\
&= 2\sum_{r=1}^{U}\frac{\binom{Z-1}{r-1}\binom{N-1}{r-1}}{\binom{Z+N}{Z}}(r-1)_{[g_0]}(r-1)_{[g_1]} \\
&\quad+\sum_{r=1}^{U}\frac{\binom{Z-1}{r-1}\binom{N-1}{r}}{\binom{Z+N}{Z}}(r-1)_{[g_0]}(r)_{[g_1]} \\
&\quad+\sum_{r=1}^{U}\frac{\binom{Z-1}{r}\binom{N-1}{r-1}}{\binom{Z+N}{Z}}(r)_{[g_0]}(r-1)_{[g_1]},
\end{aligned}$$

where U can be taken to be either Z or N (or anything larger). The first term is, say,

$$\begin{aligned}
A\binom{Z+N}{Z} &= 2\sum_{r=1}^{U}\frac{(Z-1)(Z-2)\cdots(Z+1-r)}{(r-g_0-1)!}\,\frac{(N-1)(N-2)\cdots(N+1-r)}{(r-g_1-1)!} \\
&= 2(Z-1)_{[g_0]}(N-1)_{[g_1]}\sum_{r=1}^{U}\binom{Z-1-g_0}{r-g_0-1}\binom{N-1-g_1}{r-g_1-1}
\end{aligned}$$

obtained by multiplying by the two terms

$$1=\binom{Z-1-g_0}{r-g_0-1}\frac{(r-g_0-1)!}{(Z-1-g_0)\cdots(Z-r+1)}$$

and

$$1=\binom{N-1-g_1}{r-g_1-1}\frac{(r-g_1-1)!}{(N-1-g_1)\cdots(N-r+1)}$$

for $g_0 < r$, $g_1 < r$, and simplifying. Similarly, the second term is

$$\begin{aligned}
B\binom{Z+N}{Z} &= \sum_{r=1}^{U}\frac{(Z-1)\cdots(Z+1-r)}{(r-g_0-1)!}\,\frac{(N-1)\cdots(N-r)}{(r-g_1)!} \\
&= (Z-1)_{[g_0]}(N-1)_{[g_1]}\sum_{r=1}^{U}\binom{Z-1-g_0}{r-g_0-1}\binom{N-1-g_1}{r-g_1}
\end{aligned}$$

and, by symmetry, the third must be

$$C\binom{Z+N}{Z}=(Z-1)_{[g_0]}(N-1)_{[g_1]}\sum_{r=1}^{U}\binom{Z-1-g_0}{r-g_0}\binom{N-1-g_1}{r-g_1-1}.$$

Factoring out $(Z-1)_{[g_0]}(N-1)_{[g_1]}$ from terms A, B and C, splitting up the $2\sum_{r=1}^{U}$ term in A and then using (1.4) to combine the binomial coefficients,

$$A+B+C=\frac{(Z-1)_{[g_0]}(N-1)_{[g_1]}(D+E)}{\binom{Z+N}{Z}},$$

where

$$D=\sum_{r=1}^{U}\binom{Z-1-g_0}{r-g_0-1}\binom{N-g_1}{r-g_1},\quad E=\sum_{r=1}^{U}\binom{N-1-g_1}{r-g_1-1}\binom{Z-g_0}{r-g_0}.$$

Next, as $\binom{a}{b}=0$ for $a<b$ and with $s=r-g_0-1$ and using $U=Z$,

$$D=\sum_{r=g_0+1}^{Z}\binom{Z-1-g_0}{r-g_0-1}\binom{N-g_1}{r-g_1}=\sum_{s=0}^{Z-g_0-1}\binom{Z-1-g_0}{s}\binom{N-g_1}{s+g_0+1-g_1}$$

and, similarly, but taking $U = N$,

$$E = \sum_{s=0}^{N-g_1-1} \binom{N-1-g_1}{s} \binom{Z-g_0}{s+g_1-g_0+1}.$$

Now applying (1.55), i.e.,

$$\binom{Q+M}{k+M} = \sum_{i=0}^{M} \binom{Q}{k+i} \binom{M}{i},$$

to D and E, it follows that

$$\begin{aligned} D = \sum_{i=0}^{Z-g_0-1} \binom{Z-1-g_0}{i} \binom{N-g_1}{i+g_0+1-g_1} &= \binom{N-g_1+Z-1-g_0}{g_0+1-g_1+Z-1-g_0} \\ &= \binom{Z+N-g_0-g_1-1}{Z-g_1} \end{aligned}$$

and

$$\begin{aligned} E = \sum_{s=0}^{N-g_1-1} \binom{N-1-g_1}{s} \binom{Z-g_0}{s+g_1-g_0+1} &= \binom{Z-g_0+N-1-g_1}{g_1-g_0+1+N-1-g_1} \\ &= \binom{Z+N-g_0-g_1-1}{Z-g_1-1}. \end{aligned}$$

Finally, using (1.4),

$$D + E = \binom{Z+N-g_0-g_1}{Z-g_1}$$

and the identity is proven.

Solution to Problem 6.7:

a) Generalizing (6.30) and (6.31),

$$\Pr(X = x) = \sum_{i=0}^{x} \Pr(X_1 = i) \sum_{j=0}^{x-i} \Pr(X_2 = j) \Pr(X_3 = x - i - j)$$

and

$$\Pr(X \leq x) = \sum_{i=0}^{x} \Pr(X_1 \leq i) \sum_{j=0}^{x-i} \Pr(X_2 = j) \Pr(X_3 = x - i - j).$$

b) These are seen to be

$$\mathbb{E}[X] = \sum_{i=1}^{3} \mathbb{E}[X_i] = \frac{r_1(1-p_1)}{p_1} + \frac{r_2(1-p_2)}{p_2} + \frac{r_3(1-p_3)}{p_3} = 24$$

and

$$\mathbb{V}(X) = \sum_{i=1}^{3} \mathbb{V}(X_i) = \frac{r_1(1-p_1)}{p_1^2} + \frac{r_2(1-p_2)}{p_2^2} + \frac{r_3(1-p_3)}{p_3^2} = 69.3333.$$

c) Listing S-6.1 provides a program for scalar x values. From the Matlab prompt, run the following:

```
p=zeros(71,1);
for x=0:70, p(x+1)=negbin3(x,4,1/4,8,2/4,12,3/4); end
plot(0:70, p), hold on
plot(0:70, normpdf(0:70,24,sqrt(69.3333)),'r--'), hold off
```

```
function [pdf,cdf] = negbin3 (x,r1,p1,r2,p2,r3,p3)
global cdfstart
if r1==0
  j=0:x;
  pdf = sum( negbinpdf(j,r2,p2) .* negbinpdf(x-j,r3,p3) );
  if nargout>1
    cdf = sum( cumsum(negbinpdf(j,r2,p2)) .* negbinpdf(x-j,r3,p3) );
  end
else
  cdfstart=0; pdf=0; cdf=0;
  for i=0:x
    j=0:(x-i);
    t=sum( negbinpdf(j,r2,p2) .* negbinpdf(x-i-j,r3,p3) );
    pdf=pdf+t*negbinpdf(i,r1,p1);
    if nargout>1
      cdf=cdf+t*negbincdf(i,r1,p1);
    end
  end
end

function den = negbinpdf(x,r,p)
  % den=c(r+x-1,x) .* p.^r .* (1-p).^x;
  d=gammaln(r+x) - gammaln(x+1) - gammaln(r) + r*log(p) + x * log(1-p);
  den=exp(d);

function cdf = negbincdf(x,r,p)
  global cdfstart
  start=cdfstart;       % start=sum(negbinpdf(0:(x-1),r,p));
  cdf=start + negbinpdf(x,r,p);
  cdfstart = cdf;       % prepare for next call to negbincdf
```

Program Listing S-6.1: Computation of the pmf (and cdf if requested) of $X = X_1 + X_2 + X_3$, where $X_i \overset{\text{ind}}{\sim} \text{NBin}(r_i, p_i)$. Set r_1 to zero for computing the convolution of just X_2 and X_3.

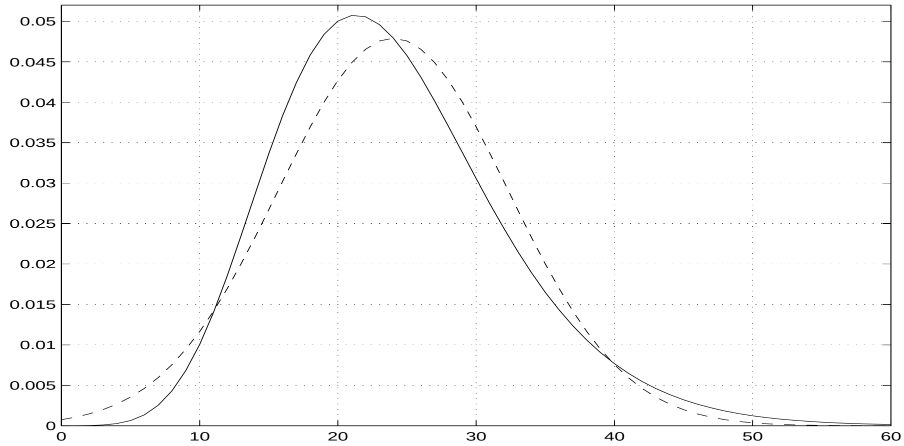


Figure S-6.1: Exact density of X (solid) and normal approximation (dashed)

The resulting graph is shown in Figure S-6.1. Clearly, X is poorly approximated as a normal distribution.

Solution to Problem 6.8:

— First note that $S_1 \sim \text{Bin}(20, p_1)$ and $S_2 \sim \text{Bin}(20, p_2)$.

a) $\Pr(S_1 > S_2) = g(20)$, where

$$g(n) := \sum_{i=1}^{n} \Pr(S_1 = i) \cdot \Pr(S_2 < i)$$
$$= \sum_{i=1}^{n} \left[\binom{n}{i} p_1^i (1 - p_1)^{n-i} \sum_{j=0}^{i-1} \binom{n}{j} p_2^j (1 - p_2)^{n-j} \right].$$

b) Similarly, $\Pr(S_1 \geq 2S_2) = h(20)$, where

$$h(n) := \sum_{i=1}^{n} \Pr(S_1 = i) \cdot \Pr\left(S_2 < \frac{i}{2}\right)$$
$$= \sum_{i=1}^{n} \binom{n}{i} p_1^i (1 - p_1)^{n-i} \sum_{j=0}^{[i/2]} \binom{n}{j} p_2^j (1 - p_2)^{n-j}.$$

c) The event $\{S_1 > S_2\}$ is a Bernoulli random variable with parameter p_{20}, where $p_{20} := g(20)$, so that $\mathbb{E}[\{S_1 > S_2\}] = p_{20}$ and $\text{Var}(\{S_1 > S_2\}) = p_{20}(1 - p_{20})$.

d) $B \sim \text{Bin}(4, p_5)$, where $p_5 := g(5)$, so $\Pr(B > 2) = \sum_{i=3}^{4} \binom{4}{i} p_5^i (1 - p_5)^{4-i}$ and $\text{Var}(B) = 4p_5(1 - p_5)$.

e) Information is lost by only examining the B_i.

Solution to Problem 6.9:

a) 1. For $k = 1$, $\{N_{1n} = 0\}$ occurs if A is already in his seat, and $\{N_{1n} = 1\}$ if not, so that

$$\Pr(N_{1n} = 0) = \frac{1}{n}, \qquad \Pr(N_{1n} = 1) = \frac{n-1}{n}$$

and

$$\mathbb{E}[N_{1n}] = 0 \cdot \Pr(N_{1n} = 0) + 1 \cdot \Pr(N_{1n} = 1) = \frac{n-1}{n}.$$

Note that event $\{N_{1n} = 1\}$ can only occur if $n \geq 2$.

2. For $k = 2$, there can be zero, one or two people at the front. Event $\{N_{2n} = 0\}$ can happen if chairs one and two are filled with A and B, i.e., the two people which belong there, in either order, so that

$$\Pr(N_{2n} = 0) = \frac{2}{n(n-1)} = \frac{\binom{2}{2}\binom{n-2}{0}}{\binom{n}{2}}.$$

Event $\{N_{2n} = 1\}$ can happen if at the beginning, the first two seats are (in either order), AX, BX, where X $\notin \{\mathrm{A}, \mathrm{B}\}$. Each of these occur with the same probability; thus

$$\Pr(N_{2n} = 1) = 4\frac{1}{n}\frac{n-2}{n-1}.$$

Lastly, event $\{N_{2n} = 2\}$ happens if chairs one and two consist of XY, where X, Y $\notin \{\mathrm{A}, \mathrm{B}\}$, with probability

$$\Pr(N_{2n} = 2) = \frac{n-2}{n}\frac{n-3}{n-1} = \frac{\binom{2}{0}\binom{n-2}{2}}{\binom{n}{2}}.$$

Note that event $\{N_{2n} = 2\}$ can occur only if $n \geq 4$. Assuming this,

$$\begin{aligned} \mathbb{E}[N_{2n}] &= 1 \cdot \frac{4}{n}\frac{n-2}{n-1} + 2 \cdot \frac{n-2}{n}\frac{n-3}{n-1} \\ &= 2\frac{n-2}{n}, \qquad n \geq 4. \end{aligned}$$

3. Now for $k = 3$. It is clear that event $\{N_{2n} = 0\}$ can happen if chairs one, two and three are filled with the three people who belong there (A,B and C), in either order, i.e.,

$$\Pr(N_{3n} = 0) = \frac{\binom{3}{3}\binom{n-3}{0}}{\binom{n}{3}} = \frac{6}{n(n-1)(n-2)}.$$

Similarly, with X $\notin \{\mathrm{A}, \mathrm{B}, \mathrm{C}\}$, event $\{N_{2n} = 1\}$ occurs when one of the events, in any of their 3! orders, $\{\mathrm{ABX}\}, \{\mathrm{ACX}\}$ and $\{\mathrm{BCX}\}$, so that

$$\Pr(N_{3n} = 1) = \frac{\binom{3}{2}\binom{n-3}{1}}{\binom{n}{3}}.$$

Likewise, with X,Y $\notin \{\mathrm{A}, \mathrm{B}, \mathrm{C}\}$, event $\{N_{2n} = 2\}$ occurs when one of the events, in any of their 3! orders, $\{\mathrm{AXY}\}, \{\mathrm{BXY}\}$ and $\{\mathrm{CXY}\}$, so that

$$\Pr(N_{3n} = 2) = \frac{\binom{3}{1}\binom{n-3}{2}}{\binom{n}{3}}.$$

The same logic leads to

$$\Pr(N_{3n} = 3) = \frac{\binom{3}{0}\binom{n-3}{3}}{\binom{n}{3}},$$

where event $\{N_{3n} = 3\}$ can occur only if $n \geq 6$. Assuming this,

$$\begin{aligned} \mathbb{E}[N_{3n}] &= \frac{1}{\binom{n}{3}}\left(1 \cdot \binom{3}{2}\binom{n-3}{1} + 2 \cdot \binom{3}{1}\binom{n-3}{2} + 3 \cdot \binom{3}{0}\binom{n-3}{3}\right) \\ &= 3\frac{n-3}{n}, \qquad n \geq 6, \end{aligned}$$

after simplifying.

b) Based on the above results, one is behooved to generalize this to $\Pr(N_{kn} = j)$ in the natural way. Obviously, $0 \leq j \leq k$, and a bit of thought and trial and error reveals that, for n even, j can be at most $n/2$, and for n odd, j can be at most $(n-1)/2$. Thus, with $\lfloor n/2 \rfloor$ the floor function,

$$\Pr(N_{kn} = j) = \frac{\binom{k}{k-j}\binom{n-k}{j}}{\binom{n}{k}} \mathbb{I}_{\{0,1,\ldots,\min(k,\lfloor n/2 \rfloor)\}}(j), \qquad 1 \leq k < n. \tag{S-6.3}$$

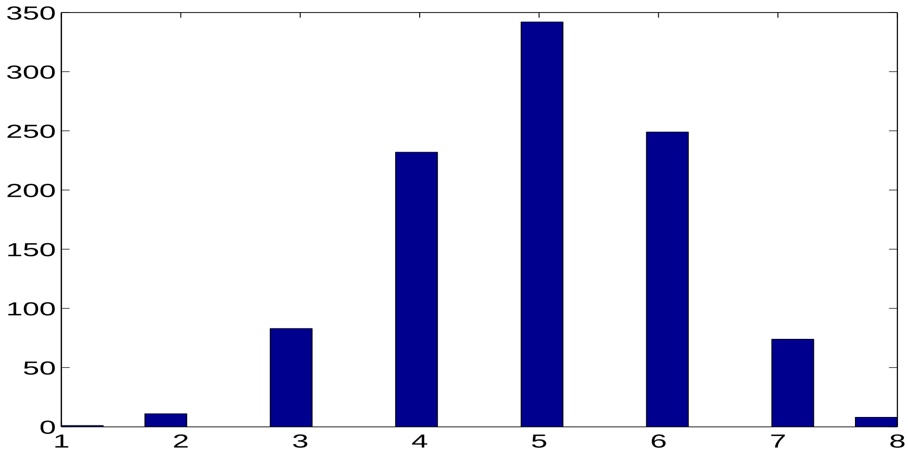


Figure S-6.2: Mass function of $N_{10,20}$ determined via simulation

That this sums to one was shown in (1.33).

Note that (S-6.3) is the pmf of a hypergeometric distribution; in particular, with the denominator in (S-6.3) written as $\binom{n}{n-k}$ instead of $\binom{n}{k}$, we see that

$$N_{kn} \sim \text{HGeo}\,(j; k, n-k, n-k)\,. \tag{S-6.4}$$

It thus follows directly from (6.21) that

$$\mathbb{E}\,[N_{kn}] = \frac{(n-k)\,k}{k+(n-k)} = k\frac{n-k}{n}.$$

c) Having worked *hard* for the previous answer and learned a bit, we can now try to work *smart*: Result (S-6.4) could be arrived at immediately by noticing that the probability that j students are standing in front after k names have been called equals the probability that, out of the $n-k$ students whose names have *not* been called, j have been sitting in one of the first k chairs.

d) See the program in Listing S-6.2. Calling `V=classroom(20,1000);` simulates the "classroom" of 20 people 1,000 times, from which any function of interest can be estimated. For example, `hist(V(:,10),20)` produces the mass function of $N_{10,20}$, as is plotted in Figure S-6.2, which can be compared to the exact values given by (S-6.3).

```
function V=classroom(n,sim,kmax)
if nargin<3, kmax=n-1; end
V=-9999*ones(sim,kmax);
for s=1:sim
  class = permvec(n); % build the class
  front=[];
  for i=1:kmax % go through the seats
    if class(i) ~= i % Person i is not in the ith seat

      if class(i)>0
        % someone is sitting there, and it is the wrong person
        front=union(front,class(i)); % put him in front of the class
        class(i)=0; % his seat is now empty
      end

      % Now the ith seat is definitely empty.
      % Let's fill it with the ith person.
      if ismember(i,front) % is the correct person in the front?
        front=setdiff(front,i); % take him out of the front...
      else % He is in some seat between the (i+1)th and nth
        loc=find(class==i); % find him
        class(loc)=0; % take him out of his seat...
      end

      class(i)=i; % ...and put him where he belongs.
    end
    V(s,i)=length(front);
  end
end

function y = permvec(N)
x=1:N; y=zeros(N,1);
for i=1:N
  p = unidrnd(N+1-i);  y(i) = x(p);  x=[x(1:(p-1)) x((p+1):(N-i+1))];
end
```

Program Listing S-6.2: Simulates the "classroom", whereby n students were seated randomly and the instructor goes through the rows, seating people alphabetically; those who are in the wrong seat go to the front of the class. For the s^{th} classroom simulation and at the i^{th} desk, $i = 1, \ldots, k_{\max}$, $V(s, i)$ contains the number of students waiting at the front of the classroom to be seated.

Solutions to Chapter 7: Continuous Univariate Random Variables

Solution to Problem 7.1:

a) We want

$$\mathbb{E}[T] = \int_{-\infty}^{\infty} t f_T(t;n)\, dt = K_n \int_{-\infty}^{\infty} t \left(1 + \frac{t^2}{n}\right)^{-k} dt,$$

where $k = (n+1)/2$. Let's ignore K_n for now, and just look at the integral. Split it at zero (because we have t^2 in the integrand) and let $u = t^2$. Then, for $t < 0$, the solution is $t = -\sqrt{u}$, with $dt = -\frac{1}{2}u^{-1/2}du$. Also, when $t = -\infty$ (the lower bound in the integral) $u = \infty$, and when $t = 0$ (the upper bound), $u = 0$. So,

$$\begin{aligned}
\int_{-\infty}^{0} t \left(1 + \frac{t^2}{n}\right)^{-k} dt &= \int_{+\infty}^{0} (-\sqrt{u}) \left(1 + \frac{u}{n}\right)^{-k} \left(-\frac{1}{2}u^{-1/2}du\right) \\
&= \frac{1}{2}\int_{\infty}^{0} \left(1 + \frac{u}{n}\right)^{-k} du \\
&= -\frac{1}{2}\int_{0}^{\infty} \left(1 + \frac{u}{n}\right)^{-k} du
\end{aligned}$$

Similarly, for $t > 0$, the solution is $t = +\sqrt{u}$, with $dt = \frac{1}{2}u^{-1/2}du$, and the integral is

$$\begin{aligned}
\int_{0}^{\infty} t \left(1 + \frac{t^2}{n}\right)^{-k} dt &= \int_{0}^{\infty} u^{1/2} \left(1 + \frac{u}{n}\right)^{-k} \frac{1}{2}u^{-1/2}du \\
&= \frac{1}{2}\int_{0}^{\infty} \left(1 + \frac{u}{n}\right)^{-k} du.
\end{aligned}$$

Thus, adding the two pieces, we get that $\mathbb{E}[T] = 0$, **if the integral**

$$I = \int_{0}^{\infty} \left(1 + \frac{u}{n}\right)^{-k} du$$

exists. Now, let $v = 1 + u/n$, so that $u = n(v-1)$ and $du = n dv$, and

$$I = n\int_{1}^{\infty} v^{-k} dv = \frac{n}{1-k}\, v^{1-k}\Big|_{v=1}^{\infty} = \begin{cases} \dfrac{n}{k-1}, & \text{if} \quad k > 1, \\ \infty, & \text{if} \quad k \le 1. \end{cases}$$

Thus, $\mathbb{E}[T] = 0$ if $k > 1$, which is the same as $k = (n+1)/2 > 1$, or $n > 1$.

b) First observe that, because of symmetry, the density of $|T|$ is just

$$f_{|T|}(t) = 2 f_T(t;n)\, \mathbb{I}_{(0,\infty)}(t).$$

Let

$$u = \frac{t^2/n}{1+t^2/n}, \quad t = +\sqrt{n\frac{u}{1-u}}, \quad dt = \frac{n^{1/2}}{2}u^{-1/2}(1-u)^{-3/2}\,du,$$

so that, after some simplifying,

$$\begin{aligned}
\mathbb{E}\big[\,|T|^k\,\big] &= 2\frac{n^{-1/2}}{B\left(\frac{n}{2},\frac{1}{2}\right)}\int_0^\infty t^k\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}dt \\
&= \frac{n^{k/2}}{B\left(\frac{n}{2},\frac{1}{2}\right)}\int_0^1 u^{(k-1)/2}(1-u)^{(n-k-2)/2}\,du \\
&= \frac{n^{k/2}}{B\left(\frac{n}{2},\frac{1}{2}\right)}B\left(\frac{k+1}{2},\frac{n-k}{2}\right).
\end{aligned}$$

This simplifies further to

$$\mathbb{E}\big[\,|T|^k\,\big] = n^{k/2}\frac{\Gamma\left(\frac{k+1}{2}\right)\Gamma\left(\frac{n-k}{2}\right)}{\Gamma\left(\frac{k+1}{2}+\frac{n-k}{2}\right)}\frac{\Gamma\left(\frac{n}{2}+\frac{1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)} = n^{k/2}\frac{\Gamma\left(\frac{k+1}{2}\right)\Gamma\left(\frac{n-k}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)}. \tag{S-7.1}$$

c) From the argument of the second gamma term in the numerator of (S-7.1), we see that $\mathbb{E}\big[\,|T|^k\,\big]$ exists only if $n > k$.

d) For $k = 2$, we have, for $n > 2$, that $\mathbb{E}\left[|T|^2\right] = \mathbb{E}\left[T^2\right]$, and $\mathbb{E}\left[T^2\right] = \mathbb{V}(T)$ because $\mathbb{E}[T] = 0$. Thus, from (S-7.1),

$$\mathbb{V}(T) = n\frac{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\;\Gamma\left(\frac{n}{2}\right)} = n\frac{\frac{1}{2}\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)}\frac{\Gamma\left(\frac{n-2}{2}\right)}{\left(\frac{n}{2}-1\right)\Gamma\left(\frac{n}{2}-1\right)} = \frac{n}{2}\frac{1}{\frac{n}{2}-1} = \frac{n}{n-2}.$$

e) Use the substitution $u = 1+x^2/n$ (so that $x = -n^{1/2}(u-1)^{1/2}$ and $dx = -n^{1/2}(1/2)(u-1)^{-1/2}\,du$), followed by $n = (u-1)/u$ (so that $u = (1-n)^{-1}$ and $du = (1-n)^{-2}\,dn$) to get

$$\begin{aligned}
F_T(t) &= \frac{n^{-1/2}}{B(n/2,1/2)}\int_{-\infty}^t \left(1+x^2/n\right)^{-(n+1)/2}dx \\
&= \frac{1}{2B(n/2,1/2)}\int_{1+t^2/n}^\infty u^{-(n+1)/2}(u-1)^{-1/2}\,du \\
&= \frac{1}{2B(n/2,1/2)}\int_{1-L}^1 n^{-1/2}(1-n)^{(n-2)/2}\,dn,
\end{aligned}$$

where

$$1-L = \frac{t^2/n}{1+t^2/n} = \frac{t^2}{n+t^2} = 1-\frac{n}{n+t^2}.$$

Thus, as

$$\int_g^1 n^{a-1}(1-n)^{b-1}\,dn = \int_0^{1-g} y^{b-1}(1-y)^{a-1}\,dy,$$

we have

$$F_T(t) = \frac{1}{2B(n/2,1/2)}\int_0^L y^{(n-2)/2}(1-y)^{-1/2}\,dy = \frac{1}{2}\bar{B}_L(n/2,1/2).$$

Solution to Problem 7.2: Letting c denote the desired quantile, we have

$$1-\alpha = \frac{2/n}{B(1,n/2)}\int_0^c\left(1+\frac{2}{n}x\right)^{-(2+n)/2}dx = 1-\left(\frac{n}{n+2c}\right)^{n/2}$$

or $c = n\left(\alpha^{-2/n}-1\right)/2$.

Solution to Problem 7.3: Using $\int u dv = uv - \int v du$ with $u = x$ and $dv = e^{-x} dx$,

$$\begin{aligned}\int_0^\infty x e^{-x} dx &= -x e^{-x}\Big|_0^\infty - \int_0^\infty -e^{-x} dx = -x e^{-x}\Big|_0^\infty - e^{-x}\Big|_0^\infty \\ &= -\lim_{x\to\infty} \frac{x}{e^x} - (-1) = 1 - \lim_{x\to\infty} \frac{1}{e^x} = 1,\end{aligned}$$

from l'Hôpital's rule. Alternatively, notice that $\int_0^\infty x e^{-x} dx = \mathbb{E}[Y]$, where $Y \sim \exp(1)$, but $\mathbb{E}[Y] = 1$.

Solution to Problem 7.4: Some trial and error shows that the substitution $u = r/(1+r)$ is advantageous. With it, $r = u/(1-u)$, $1 + r = 1/(1-u)$ and $dr = (1-u)^{-2} du$, so that

$$\begin{aligned}\int_0^1 r^{a-1} (1+r)^{-2a} dr &= \int_0^{1/2} \left(\frac{u}{1-u}\right)^{a-1} (1-u)^{2a} (1-u)^{-2} du \\ &= \int_0^{1/2} u^{a-1} (1-u)^{a-1} du = \frac{B(a,a)}{2}.\end{aligned}$$

Similarly, the expected value of R is

$$\begin{aligned}\int_0^1 r f_R(r)\, dr &= \frac{2}{B(a,a)} \int_0^1 r^a (1+r)^{-2a} dr \\ & \frac{2}{B(a,a)} \int_0^{1/2} \left(\frac{u}{1-u}\right)^a (1-u)^{2a} (1-u)^{-2} du \\ &= \frac{2}{B(a,a)} \int_0^{1/2} u^a (1-u)^{a-2} du = 2\frac{B_{1/2}(a+1, a-1)}{B(a,a)},\end{aligned}$$

as was determined in Example 7.10.

Solution to Problem 7.5: Using the substitution $c = \nu x^2$, then $dc/dx = 2\nu x$ and, from (7.65),

$$\begin{aligned}f_X(x) &= \left|\frac{dc}{dx}\right| f_C(c) = 2\nu x \cdot \frac{2^{-\nu/2}}{\Gamma(\nu/2)} \left(\nu x^2\right)^{\nu/2-1} e^{-(\nu x^2)/2} \mathbb{I}_{(0,\infty)}\left(\nu x^2\right) \\ &= \frac{2^{-\nu/2+1} \nu^{\nu/2}}{\Gamma(\nu/2)} x^{\nu-1} e^{-(\nu x^2)/2} \mathbb{I}_{(0,\infty)}(x).\end{aligned}$$

For $\nu = 1$,

$$f_X(x) = \frac{\sqrt{2}}{\sqrt{\pi}} \exp\left\{-\frac{1}{2}x^2\right\} \mathbb{I}_{(0,\infty)}(x),$$

with expected value

$$\mathbb{E}[X] = \frac{\sqrt{2}}{\sqrt{\pi}} \int_0^\infty x \exp\left\{-\frac{1}{2}x^2\right\} dx = \frac{\sqrt{2}}{\sqrt{\pi}} \left(-e^{-\frac{1}{2}x^2}\right)\Big|_0^\infty = \frac{\sqrt{2}}{\sqrt{\pi}},$$

which is, of course, the same result as given in (7.39). For the second moment, substitute $u = x^2/2$ so that $x = +\sqrt{2u}$ (because $x > 0$), $dx = (2u)^{-1/2} du$ and

$$\mathbb{E}\left[X^2\right] = \frac{2}{\sqrt{\pi}} \int_0^\infty u^{1/2} \exp\{-u\}\, du = \frac{2}{\sqrt{\pi}} \Gamma(3/2) = 1,$$

recalling that $\Gamma(3/2) = \sqrt{\pi}/2$. Thus, $\mathbb{V}(X) = 1 - 2/\pi \approx 0.36338$.

Solution to Problem 7.6: Similar to Example 7.14,

$$\begin{aligned}F_Y(y) &= \Pr(Y \le y) = \Pr(-\sqrt{y} \le X \le \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}) \\ &= \Phi\left(\frac{\sqrt{y} - \mu}{\sigma}\right) - \Phi\left(\frac{-\sqrt{y} - \mu}{\sigma}\right).\end{aligned}$$

Differentiating,

$$\begin{aligned} f_Y(y) &= \frac{1}{2\sigma\sqrt{y}}\left(\phi\left(\frac{\sqrt{y}-\mu}{\sigma}\right)+\phi\left(\frac{-\sqrt{y}-\mu}{\sigma}\right)\right) \\ &= \frac{1}{2\sigma\sqrt{y}}\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{1}{2}\left(\frac{\sqrt{y}-\mu}{\sigma}\right)^2\right\}+\exp\left\{-\frac{1}{2}\left(\frac{-\sqrt{y}-\mu}{\sigma}\right)^2\right\}\right) \end{aligned}$$

for $y > 0$. This can also be written as

$$f_Y(y) = \frac{1}{2\sigma\sqrt{y}}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{y+\mu^2}{2\sigma^2}\right)\left(\exp\left(\frac{\mu\sqrt{y}}{\sigma^2}\right)+\exp\left(-\frac{\mu\sqrt{y}}{\sigma^2}\right)\right).$$

The expected value appears difficult to calculate directly, but there is no need: From (4.49), $\mathbb{E}[Y] = \mathbb{E}[X^2] = \sigma^2+\mu^2$. For the variance, with $k = \mathbb{E}[X^2] = \sigma^2+\mu^2$,

$$\mathbb{V}(Y) = \mathbb{V}(X^2) = \mathbb{E}\left[\left(X^2-k\right)^2\right] = \mathbb{E}\left[X^4\right]-k^2$$

and, from (4.49) and Example 7.3,

$$\mathbb{E}\left[X^4\right] = \mu_4+4\mu_3\mu+6\mu_2\mu^2+\mu^4 = 3\sigma^4+6\sigma^2\mu^2+\mu^4$$

so that

$$\begin{aligned} \mathbb{V}(Y) &= 3\sigma^4+6\sigma^2\mu^2+\mu^4-\left(\sigma^2+\mu^2\right)^2 \\ &= 2\sigma^2\left(\sigma^2+2\mu^2\right). \end{aligned}$$

Solution to Problem 7.7: With $y = ax^b$, $x = (y/a)^{1/b}$ and $dx = (ab)^{-1}(y/a)^{1/b-1}\,dy$ so that, from (7.65),

$$\begin{aligned} f_Y(y) &= f_X(x)\left|\frac{dx}{dy}\right| = \frac{n_1/n_2}{ab\,B\left(\frac{n_1}{2},\frac{n_2}{2}\right)}\frac{\left(\frac{n_1}{n_2}(y/a)^{1/b}\right)^{n_1/2-1}}{\left(1+\frac{n_1}{n_2}(y/a)^{1/b}\right)^{(n_1+n_2)/2}}\left(\frac{y}{a}\right)^{1/b-1}\mathbb{I}_{(0,\infty)}(y) \\ &= \frac{a^{n_2/(2b)}}{b}\frac{(n_2/n_1)^{n_2/2}}{B(n_1/2,n_2/2)}\frac{y^{(n_1-2b)/(2b)}}{\left(a^{1/b}n_2/n_1+y^{1/b}\right)^{(n_1+n_2)/2}}\mathbb{I}_{(0,\infty)}(y) \\ &= \frac{(n_1/n_2)^{n_1/2}}{ab\,B(n_1/2,n_2/2)}\frac{(y/a)^{n_1/(2b)-1}}{\left(1+(n_1/n_2)(y/a)^{1/b}\right)^{(n_1+n_2)/2}}\mathbb{I}_{(0,\infty)}(y), \end{aligned}$$

after some simplification. The cdf is easily seen to be

$$F_Y(y) = F_X\left((y/a)^{1/b}\right) = \bar{B}_y\left(\frac{n_1}{2},\frac{n_2}{2}\right), \quad y = \frac{n_1(y/a)^{1/b}}{n_2+n_1(y/a)^{1/b}}.$$

The raw moments are simply given by $\mathbb{E}[Y^r] = \mathbb{E}\left[\left(aX^b\right)^r\right] = a^r\mathbb{E}\left[X^{br}\right]$.

Solution to Problem 7.8: From (7.68) we require

$$\lim_{x\to0^+}\frac{f_Z(x)}{x^r} = \lim_{x\to0^+}\frac{\beta^\alpha}{\Gamma(\alpha)}\frac{x^{\alpha-1}e^{-\beta x}}{x^r} \propto \lim_{x\to0^+}x^{\alpha-1-r}$$

to be finite. This holds when $\alpha > 1$. Recall that this is a sufficient (but not necessary) condition, so that this does *not* show that $\mathbb{E}\left[Z^{-1}\right]$ does *not* exist for $\alpha \le 1$.

Solution to Problem 7.9:

a) For $\mathbb{E}[Z]$ with $c = \beta z$,

$$\mathbb{E}[Z] = \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} z^\alpha \exp(-\beta z)\, dz = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty \left(\frac{c}{\beta}\right)^\alpha \exp(-c) \frac{1}{\beta} dc$$
$$= \frac{1}{\beta \Gamma(\alpha)} \Gamma(\alpha+1) = \frac{\alpha}{\beta}.$$

Similarly,

$$\mathbb{E}\left[Z^{-1}\right] = \frac{\beta}{\alpha-1}.$$

Both of these could also be obtained directly from the more general expression in (7.9).

b) With $z = 1/y$ and $dz/dy = -y^{-2}$, (7.65) gives

$$f_Y(y) = f_Z(z) \cdot y^{-2} = \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{y}\right)^{\alpha-1} e^{-\frac{1}{y}\beta} \mathbb{I}_{(0,\infty)}(z) \cdot y^{-2}$$
$$= \frac{\beta^\alpha}{\Gamma(\alpha)} y^{-(\alpha+1)} e^{-\frac{\beta}{y}} \mathbb{I}_{(0,\infty)}(y). \qquad \text{(S-7.2)}$$

As $f(y;\alpha,\beta) = \beta^{-1} f(y/\beta;\alpha,1)$, β is a (genuine) scale parameter. Of course, this follows without recourse to the density, because, if $G \sim \text{Gam}(\alpha, 1)$, then $Z = G/\beta \sim \text{Gam}(\alpha,\beta)$, $G^{-1} \sim \text{IGam}(\alpha, 1)$ and $Y = 1/Z = \beta/G$ is just a scale value (β) times a scale–one inverse gamma random variable.

c) Because $Y > 0$,

$$\Pr(Y \le y) = \Pr\left(\frac{1}{Y} \ge \frac{1}{y}\right) = 1 - \Pr\left(Z \le y^{-1}\right),$$

where $Z \sim \text{Gam}(a, b)$.

d) Using (S-7.2) and the change of variable $x = \beta/y$, $y = \beta/x$, $dy = -\beta x^{-2} dx$,

$$\mu_r' = \mathbb{E}[Y^r] = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty y^{r-(\alpha+1)} e^{-\frac{\beta}{y}} dy = -\frac{\beta^\alpha}{\Gamma(\alpha)} \int_\infty^0 \left(\frac{\beta}{x}\right)^{r-(\alpha+1)} e^{-x} \beta x^{-2} dx$$
$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty \left(\frac{\beta}{x}\right)^{r-(\alpha+1)} e^{-x} \beta x^{-2} dx = \frac{\beta^r}{\Gamma(\alpha)} \int_0^\infty \left(\frac{1}{x}\right)^{r-\alpha+1} e^{-x} dx$$
$$= \frac{\beta^r}{\Gamma(\alpha)} \int_0^\infty x^{(\alpha-r)-1} e^{-x} dx = \frac{\Gamma(\alpha-r)}{\Gamma(\alpha)} \beta^r.$$

For $r = 1$,

$$\mathbb{E}[Y] = \beta \frac{\Gamma(\alpha-1)}{\Gamma(\alpha)} = \frac{\beta}{\alpha-1},$$

which is, of course, just $\mathbb{E}\left[Z^{-1}\right]$ as found previously. From (7.9),

$$\mathbb{E}\left[Z^{-2}\right] = \frac{\beta^2}{(\alpha-1)(\alpha-2)}$$

so that, with $Y = 1/Z$,

$$\mathbb{V}(Y) = \mathbb{E}\left[Y^2\right] - (\mathbb{E}[Y])^2 = \frac{\beta^2}{(\alpha-1)(\alpha-2)} - \left(\frac{\beta}{\alpha-1}\right)^2$$
$$= \frac{\beta^2}{\alpha-1}\left(\frac{1}{\alpha-2} - \frac{1}{\alpha-1}\right) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}. \qquad \text{(S-7.3)}$$

e) Clearly, $\mathbb{E}[Z]$ and $\mathbb{E}\left[Z^{-1}\right]$ are not reciprocals, but

$$\frac{\alpha}{\beta} - \frac{\alpha-1}{\beta} = \beta^{-1} \qquad \text{(S-7.4)}$$

and

$$\frac{\beta}{\alpha-1}-\frac{\beta}{\alpha}=\frac{\beta}{\alpha\left(\alpha-1\right)} \tag{S-7.5}$$

so that, from the first, β needs to be large and, from the second, α^2 needs to be large, but such that β/α^2 goes to zero. This is satisfied, for example, with $\alpha = c\beta$, $c > 0$, and then taking the limit as $\beta \to \infty$. In particular, (S-7.4) clearly goes to zero as $\beta \to \infty$, while (S-7.5) gives

$$\lim_{\beta\to\infty}\frac{\beta}{\alpha\left(\alpha-1\right)}=\lim_{\beta\to\infty}\frac{1}{c\left(c\beta-1\right)}=0.$$

Notice that $\mathbb{E}\left[Z\right]=\alpha/\beta=c\beta/\beta=c$ stays constant, but $\mathbb{V}\left(Z\right)=\alpha/\beta^2=c/\beta\to 0$. Thus, as α and β increase such that $\alpha = c\beta$, the density piles up onto the point c. Clearly then, the density of its reciprocal will pile up onto $1/c$, with expected value $1/c$ and zero variance. To see the latter, from (S-7.3),

$$\lim_{\substack{\beta\to\infty\\ \alpha=c\beta}}\mathbb{V}\left(Y\right)=\lim_{\beta\to\infty}\frac{\beta^2}{\left(c\beta-1\right)^2\left(c\beta-2\right)}=0.$$

Solution to Problem 7.10: With $y = x - 1$,

$$I=\int_{-1}^{1}\left(y+1\right)e^{-y^2}dy=\int_{-1}^{1}ye^{-y^2}dy+\int_{-1}^{1}e^{-y^2}dy=0+\int_{-1}^{1}e^{-y^2}dy$$

and, with $\sigma^2 = 1/2$,

$$\int_{-1}^{1}e^{-y^2}dy=\sqrt{2\pi}\sigma\int_{-1}^{1}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{y^2}{2\sigma^2}}dy=\sqrt{2\pi}\sigma\Pr\left(-1\le X\le 1\right)$$

where $X \sim \mathrm{N}\left(0,\sigma^2\right)$. But

$$\Pr\left(-1\le X\le 1\right)=\Pr\left(\frac{-1}{1/\sqrt{2}}\le Z\le\frac{1}{1/\sqrt{2}}\right)=1-2\Phi\left(-\sqrt{2}\right),$$

where $Z \sim \mathrm{N}(0,1)$, i.e., $I=\sqrt{2\pi}\sigma\left(1-2\Phi\left(-\sqrt{2}\right)\right)=1.4936483$.

Solution to Problem 7.11: We recognize I to be $\mathbb{E}\left[X^2\right]$, where $X\sim\mathrm{N}\left(0,\sigma^2\right)$, $\sigma=1/\sqrt{2}$, without the constant term. As $\mathbb{E}\left[X^2\right]=\mu^2+\sigma^2=1/2$,

$$\frac{1}{2}=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}x^2\exp\left(-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right)dx=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}x^2e^{-x^2}dx,$$

so that $I=\sqrt{\pi}/2$.

Solution to Problem 7.12:

a) Necessity means that, if $\mathbb{E}\left[X\right]$ exists, then

$$\lim_{x\to\infty}x\left(1-F_X\left(x\right)\right)=0.$$

If $\mathbb{E}\left[X\right]=\int_0^\infty t\,f_X\left(t\right)dt$ exists, then

$$\lim_{x\to\infty}\int_x^\infty t\,f_X\left(t\right)dt=0.$$

So, as $\lim_{x\to\infty}x\left(1-F_X\left(x\right)\right)\ge 0$ and

$$\lim_{x\to\infty}x\left(1-F_X\left(x\right)\right)=\lim_{x\to\infty}x\int_x^\infty f_X\left(t\right)dt\le\lim_{x\to\infty}\int_x^\infty t\,f_X\left(t\right)dt=0,$$

it follows that $\lim_{x\to\infty} x\left(1 - F_X(x)\right) = 0$.
To show that the expectation of a Cauchy random variable does not exist, note that the theorem is valid for nonnegative r.v.s, so, we must first show that we apply it to $\mathbb{E}\big[|X|\big]$.
Observe that, if X is a random variable with density f_X symmetric about zero, then $f_X(-x) = f_X(x)$ and, substituting $u = -x$,

$$\mathbb{E}[X] = \int_{-\infty}^{0} x f_X(x)\,\mathrm{d}x + \int_{0}^{\infty} x f_X(x)\,\mathrm{d}x = -\int_{0}^{\infty} u f_X(u)\,\mathrm{d}u + \int_{0}^{\infty} x f_X(x)\,\mathrm{d}x = 0,$$

if $\int_0^\infty x f_X(x)\mathrm{d}x$ exists. If it does, then, as $\mathbb{E}[g(X)] = \int g(x) f(x)\mathrm{d}x$,

$$\begin{aligned}
\mathbb{E}[|X|] &= \int_{-\infty}^{0} |x| f_X(x)\,\mathrm{d}x + \int_{0}^{\infty} |x| f_X(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{0} (-x) f_X(x)\,\mathrm{d}x + \int_{0}^{\infty} x f_X(x)\,\mathrm{d}x \\
&= -\int_{-\infty}^{0} x f_X(x)\,\mathrm{d}x + \int_{0}^{\infty} x f_X(x)\,\mathrm{d}x = 2\int_{0}^{\infty} x f_X(x)\,\mathrm{d}x,
\end{aligned}$$

where $\int_{-\infty}^{0} x f_X(x)\mathrm{d}x = -\int_0^\infty u f_X(u)\mathrm{d}u$ from above.
The latter result can also be seen as follows. If $Z = |X|$ and f_X is symmetric about zero, then

$$\begin{aligned}
F_Z(z) &= \Pr(Z \le z) = \Pr(|X| \le z) = \Pr(-z \le X \le z) \\
&= F_X(z) - F_X(-z) = [1 - F_X(-z)] - F_X(-z) \\
&= 1 - 2F_X(-z)
\end{aligned}$$

and

$$f_Z(z) = \frac{\mathrm{d}F_Z(z)}{\mathrm{d}z} = -(-1)2f_X(-z) = 2f_X(z),$$

so that $\mathbb{E}[|X|] = 2\int_0^\infty x f_X(x)\,\mathrm{d}x$.
We see that, if $f_X(-x) = f_X(x)$ and $\mathbb{E}[X]$ exists, then $\mathbb{E}[|X|]$ exists. The contrapositive then implies that, if $\mathbb{E}[|X|]$ does not exist, then $\mathbb{E}[X]$ does not exist.
If X is the absolute value of a standard Cauchy random variable, then

$$f_X(x) = \frac{2}{\pi}\frac{1}{1+x^2} \qquad \text{and} \qquad F_X(x) = \frac{2}{\pi}\arctan(x).$$

As $\lim_{x\to\infty}(1 - F_X(x)) = 0$, l'Hôpital's rule gives

$$\lim_{x\to\infty} x\left(1 - F_X(x)\right) = \lim_{x\to\infty} \frac{\left(1 - \frac{2}{\pi}\arctan(x)\right)}{1/x} = \lim_{x\to\infty} \frac{-\frac{2}{\pi}\frac{1}{1+x^2}}{-x^{-2}} = \frac{2}{\pi},$$

which is nonzero.
The necessity of the condition implies that $\mathbb{E}[|X|]$, and thus, $\mathbb{E}[X]$, does not exist.

b) Let $u = 1 - F_X(x)$ and $dv = dx$, so that

$$\begin{aligned}
I &= \int_0^\infty \left(1 - F_X(x)\right)\cdot 1\ dx \\
&= uv|_0^\infty - \int_0^\infty v du = x\left(1 - F_X(x)\right)|_0^\infty - \int_0^\infty (-1)\, x F_X'(x)\, dx \\
&= \lim_{x\to\infty} x\left(1 - F_X(x)\right) + \int_0^\infty x f_X(x)\, dx \\
&= \mathbb{E}[X],
\end{aligned}$$

provided $\lim_{x\to\infty} x\left(1 - F_X(x)\right) = 0$, which is a necessary condition for the existence of the first moment, as shown in the previous question.

Solution to Problem 7.13:

a) From the sketch of the two dimensional integration space depicted in Figures (S-7.1) and (S-7.2) and assuming the order of integration can be interchanged, it follows that

$$\begin{aligned}\int_0^\infty (1 - F_X(x))\,dx &= \int_0^\infty \int_x^\infty f_X(t)\,dt\,dx \\ &= \int_0^\infty \left(\int_0^t dx\right) f_X(t)\,dt = \int_0^\infty t f_X(t)\,dt,\end{aligned}$$

and the last expression is just $\mathbb{E}[X]$.

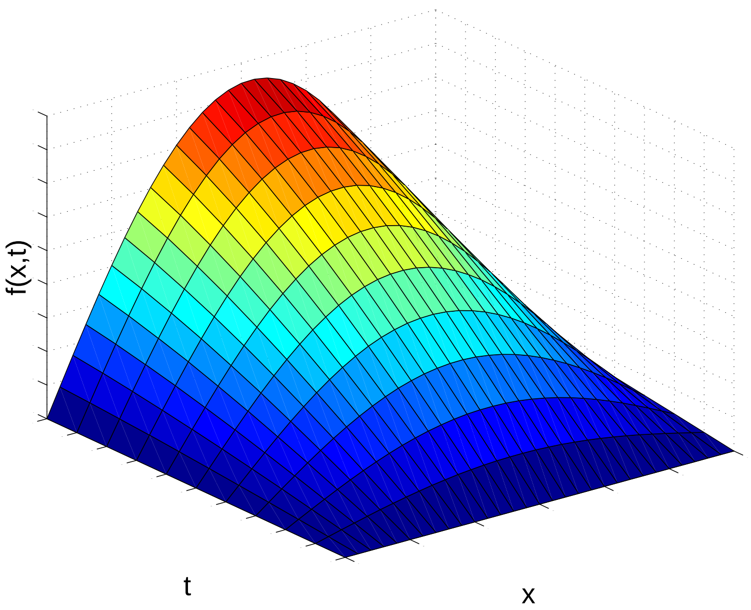


Figure S-7.1: Some function $f(x,t)$ for illustration. Notice that $f_X(t)$ is not a function of x (and $\int_x^\infty f_X(t)\,dt$ is not a function of t) but, just for integration purposes, we can think of f as a general (continuous) function of both x and t.

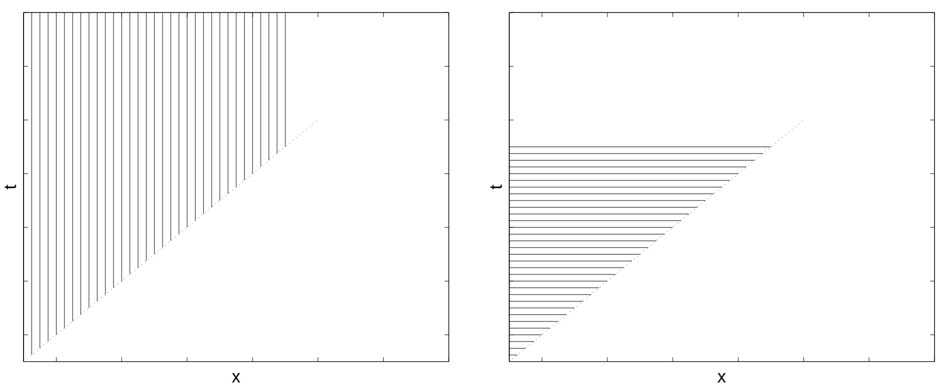


Figure S-7.2: Left panel shows the x, t surface over which we integrate f, whereby the outer integral is w.r.t. x and, thus, for a fixed x (signified by the vertical lines), the inner integral over t goes from x to ∞. The right panel illustrates the same surface but with the outer integral w.r.t. t, so that, for a fixed t (horizontal lines), the inner integral over x goes from 0 to t.

b) Similar to the previous derivation,

$$\begin{aligned}\mathbb{E}[X^n] &= \int_0^\infty t^n f_X(t)\,dt = \int_0^\infty \left(\int_0^t n x^{n-1} dx\right) f_X(t)\,dt \\ &= \int_0^\infty n x^{n-1} \int_x^\infty f_X(t)\,dt\,dx = \int_0^\infty n x^{n-1}(1 - F_X(x))\,dx.\end{aligned}$$

c) For $x < 0$,

$$\begin{aligned}\int_{-\infty}^{0} F_X(x)\,dx &= \int_0^{\infty} F_X(-x)\,dx = \int_0^{\infty}\int_{-\infty}^{-x} f_X(t)\,dtdx \\ &= \int_{-\infty}^{0}\int_0^{-t} f_X(t)\,dxdt \\ &= \int_{-\infty}^{0} f_X(t)(-t)\,dt = -\int_{-\infty}^{0} t f_X(t)\,dt.\end{aligned}$$

Combining this with the derivation for $x > 0$ used for (7.69) yields

$$-\int_{-\infty}^{0} F_X(x)\,dx + \int_0^{\infty}(1 - F_X(x))\,dx = \int_{-\infty}^{0} t f_X(t)\,dt + \int_0^{\infty} t f_X(t)\,dt = \mathbb{E}[X].$$

Solution to Problem 7.14:

a) We require that value of x such that $f_X(x)$ attains its maximum, i.e., where $\frac{\partial}{\partial x} f_X(x) = 0$. Differentiating gives

$$\frac{\partial}{\partial x} x^{a-1}(1-x)^{b-1} = -(b-1)x^{a-1}(1-x)^{b-2} + (a-1)x^{a-2}(1-x)^{b-1} = 0$$

or $(b-1)x^{a-1}(1-x)^{b-2} = (a-1)x^{a-2}(1-x)^{b-1}$ or $(b-1)x = (a-1)(1-x)$, or

$$\text{mode}(X) = \frac{a-1}{a+b-2}.$$

b) With $a = 3$ and $b = 2$ we want $\frac{(a+b-1)!}{(a-1)!(b-1)!}\int_0^x t^{a-1}(1-t)^{b-1}\,dt = 4x^3 - 3x^4 = 0.5$, or $x = 0.614272$.

c) Symmetry implies $a = b$, yielding $\mathbb{V}(X) = \frac{1}{4\times(2a+1)} = 0.2^2$, or $a = b = 2.625$.

d) We have

$$\frac{5}{8} = \frac{a-1}{a+b-2} = \frac{a-1}{a-1+b-1},$$

which can be simplified to $3(a-1) = 5(b-1)$. Because both, a and b are integers, they can be expressed as $(a-1) = 5n$ and $(b-1) = 3n$ for some integer $n \in \mathbb{N}$, so the set of all possible combinations (a, b) can be written as

$$\{(a,b) : a = 5n+1, b = 3n+1, n \in \mathbb{N}\}.$$

Thus, we have the approximate relation

$$\sigma_X(n) := \sqrt{\frac{ab}{(a+b+1)(a+b)^2}} = \frac{1}{2}\sqrt{\frac{(1+5n)(1+3n)}{(3+8n)(1+4n)^2}} \approx 0.1.$$

Trying a few values, $\sigma_X(1) = 0.148, \sigma_X(2) = 0.112$, and $\sigma_X(3) = 0.094$, so that either $n = 2$ or $n = 3$, i.e., $\{a, b\} = \{11, 7\}$ or $(a, b) = (16, 10)$ are good choices. Note that we have the exact solution of $\sigma_X(n) = 0.1$ with $n \in \mathbb{R}_{>0}$ near $n = 2.58$.

Solution to Problem 7.15: With $u = \frac{1}{2}x^2$ and making use of the $\Gamma(\cdot)$ function,

$$\begin{aligned}\mathbb{E}|X|^p &= \frac{2}{\sqrt{2\pi}}\int_0^{\infty}(2u)^{p/2}e^{-u}(2u)^{-1/2}\,du \\ &= \frac{2^{p/2}}{\sqrt{\pi}}\int_0^{\infty} u^{\frac{p+1}{2}-1}e^{u}du = \frac{2^{p/2}}{\sqrt{\pi}}\Gamma\left(\frac{p+1}{2}\right),\end{aligned}$$

so that

$$\mathbb{E}|X|^p = \begin{cases} 1, & \text{if} \quad p = 0, \\ \sqrt{2/\pi}, & \text{if} \quad p = 1, \\ 1 & \text{if} \quad p = 2, \end{cases}$$

because $\Gamma(3/2) = \sqrt{\pi}/2$.

Solution to Problem 7.16:

a) With $u = x^p$,

$$\int_{-\infty}^{\infty} f_X(x;p)\,dx = \frac{p}{\Gamma(p^{-1})}\int_0^{\infty} e^{-x^p}dx = \frac{p}{\Gamma(p^{-1})}\frac{1}{p}\int_0^{\infty} e^{-u}u^{1/p-1}dx = 1.$$

b) The cdf for $x < 0$ is given by

$$F_X(x;p) = \frac{p}{2\Gamma(p^{-1})}\int_{-\infty}^{x}\exp\{-(-t)^p\}\,dt.$$

Use the substitution

$$u = (-t)^p, \quad t = -u^{1/p}, \quad \frac{dt}{du} = -\frac{1}{p}u^{\frac{1-p}{p}}$$

so that, recalling the definition of the incomplete gamma function and the incomplete gamma ratio from §1.5.1, the cdf $F_X(x)$ for $x < 0$ is

$$\begin{aligned}\frac{1}{2\Gamma(p^{-1})}\int_{(-x)^p}^{\infty}\exp\{-u\}\,u^{(p^{-1}-1)}du &= \frac{1}{2}\left(1-\frac{\Gamma_{(-x)^p}(p^{-1})}{\Gamma(p^{-1})}\right) \qquad \text{(S-7.6)}\\ &= \frac{1}{2}\left(1-\bar{\Gamma}_{(-x)^p}(p^{-1})\right).\end{aligned}$$

Note that, as $x \uparrow 0$, $F_X(x;p) \uparrow 1/2$, as it should, given the symmetry of the density. This symmetry also implies that

$$F_X(x) = 1 - F_X(-x),$$

from which $F_X(x)$ for $x > 0$ can be computed using (S-7.6).

Recall that, in Matlab, their "incomplete gamma function", `gammainc`, computes the incomplete gamma *ratio*. So, a program for the cdf would look like that in Listing S-7.1.

c) From (7.1),

$$f_Y(y) = \sigma^{-1}f_X(y/\sigma) = \frac{p}{2\sigma\Gamma(p^{-1})}\exp\left\{-\left|\frac{x}{\sigma}\right|^p\right\},$$

the Laplace is obtained for $p = 1$; with $p = 2$ and replacing σ by $\sqrt{2}\sigma_{\mathrm{N}}$, $Y \sim \mathrm{N}\left(0, \sigma_{\mathrm{N}}^2\right)$.

d) $\mathbb{V}(Y) = \sigma^2\mathbb{V}(X) = \sigma^2\mathbb{E}\left[X^2\right]$ (because, from symmetry, $\mathbb{E}[X] = 0$) and

$$\begin{aligned}\sigma^2\mathbb{E}\left[X^2\right] &= \frac{\sigma^2 p}{2\Gamma(p^{-1})}\int_{-\infty}^{\infty}x^2e^{-|x|^p}dx\\ &= \frac{\sigma^2 p}{2\Gamma(p^{-1})}\left[\int_{-\infty}^{0}x^2e^{-((-x)^p)}dx + \int_0^{\infty}x^2e^{-x^p}dx\right]\\ &= \frac{\sigma^2 p}{\Gamma(p^{-1})}\int_0^{\infty}x^2e^{-x^p}dx = \frac{\sigma^2}{\Gamma(p^{-1})}\int_0^{\infty}u^{2/p}e^{-u}u^{1/p-1}du\\ &= \frac{\sigma^2\Gamma(3/p)}{\Gamma(1/p)},\end{aligned}$$

where the substitution $v = -x$ easily shows that

$$\int_{-\infty}^{0}x^2e^{-((-x)^p)}dx = \int_0^{\infty}v^2e^{-(v^p)}dv$$

and the substitution $u = x^p$ gives the second to last equality.

For $p = 1$, $\mathbb{V}(Y)$ reduces to $2\sigma^2$ and, for $p = 2$ and $\sigma \leftarrow \sqrt{2}\sigma_{\mathrm{N}}$, $\mathbb{V}(Y)$ reduces to σ_{N}^2, recalling that $\Gamma(3/2) = \Gamma(1/2)/2$.

```
function cdf = gedCDF(yvec,p,mu,scale,numint);

if nargin<3, mu=0; end
if nargin<4, scale=1; end
if nargin<5, numint=0; end

k=sqrt(2); % Set k to sqrt(2), and the GED with p=2
           %   coincides with the standard normal.
           %   Set k=1, and GED with p=1 coincides with the Laplace.
scale=scale*k;

zvec = (yvec-mu)./scale;
cdf=zeros(length(zvec),1);
for i=1:length(zvec)
  z=zvec(i);
  if numint==0
    if z<=0
      t = (1/2) * ( 1 - gammainc((-z)^p, 1/p) );
    else
      t = 1- (  (1/2) * ( 1 - gammainc(z^p, 1/p)));
    end;
  else
    t = quadl(@geddenstandard,-35,z,1e-8,0,p);
  end
  cdf(i)=t;
end


function f=geddenstandard(z,p)
f = p./(2*gamma(1./p)) .* exp(-abs(z).^p);
```

Program Listing S-7.1: Computes the cdf of a location–scale GED distribution. The factor k can be used to make the GED parameterization such that it coincides with the standard Laplace ($k = 1$) or the standard normal ($k = \sqrt{2}$). The `numint` option allows the user to use numeric integration to check the results.

e) From the symmetry of the GED density, it follows from (7.65) and the examples in §7.3.2 that

$$f_Y(y) = 2 \cdot f_X(x) \left|\frac{dx}{dy}\right| \mathbb{I}_{(0,\infty)}(x).$$

However, for practice' sake, we once again confirm this, and write

$$f_Y(y) = f_X(x) \left|\frac{dx}{dy}\right| \mathbb{I}_{(0,\infty)}(x) + f_X(x) \left|\frac{dx}{dy}\right| \mathbb{I}_{(-\infty,0]}(x),$$

where f_X is the GED density in (7.73). For the first term, solving $y = |x|^p$ with $x > 0$ gives $x = y^{1/p}$, and $dx/dy = (1/p)\, y^{1/p-1}$. Solving $y = |x|^p$ with $x \leq 0$ gives $x = -\left(y^{1/p}\right)$ and $dx/dy = -(1/p)\, y^{1/p-1}$. As

$$\mathbb{I}_{(-\infty,0]}(x) = \mathbb{I}_{(-\infty,0]}\left(-\left(y^{1/p}\right)\right) = \mathbb{I}_{[0,\infty)}\left(y^{1/p}\right) = \mathbb{I}_{[0,\infty)}(y),$$

we get

$$\begin{aligned} f_Y(y) &= \left|\frac{1}{p} y^{1/p-1}\right| \frac{p}{2\Gamma(1/p)} \exp\left\{-\left|y^{1/p}\right|^p\right\} \mathbb{I}_{(0,\infty)}(y) \\ &\quad + \left|-\frac{1}{p} y^{1/p-1}\right| \frac{p}{2\Gamma(1/p)} \exp\left\{-\left|-\left(y^{1/p}\right)\right|^p\right\} \mathbb{I}_{(0,\infty)}(y) \\ &= \frac{1}{\Gamma(1/p)} y^{1/p-1} \exp\{-y\} \mathbb{I}_{(0,\infty)}(y), \end{aligned}$$

which is the gamma density with scale one and shape $1/p$.

Solution to Problem 7.17:

$$\mathbb{E}\left[e^{tZ}\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2} z^2 + tz\right\} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left(z^2 - 2tz\right)\right\} dz$$

and, by completing the square as $z^2 - 2tz + t^2 - t^2 = (t-z)^2 - t^2$,

$$\begin{aligned} \mathbb{E}\left[e^{tZ}\right] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left((t-z)^2 - t^2\right)\right\} dz \\ &= \exp\left\{\frac{t^2}{2}\right\} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(z-t)^2\right\} dz \\ &= \exp\left\{\frac{t^2}{2}\right\}, \end{aligned}$$

as the last integrand is the $\mathrm{N}(t,1)$ density. Because $X = \mu + \sigma Z$ is a (location-scale) transformation of X,

$$\mathbb{E}\left[e^{tX}\right] = \mathbb{E}\left[e^{t(\mu+\sigma Z)}\right] = \exp\{t\mu\} M_Z(t\sigma) = \exp\left\{t\mu + \frac{t^2\sigma^2}{2}\right\}.$$

Solution to Problem 7.18: Let

$$q = \frac{2tb}{(1-2t)}.$$

Then

$$
\begin{aligned}
I &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(t\,(z+b)^2 - \frac{1}{2}z^2\right) dz \\
&= \frac{1}{\sqrt{2\pi}} e^{tb^2} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(1-2t)\,z^2 + 2tzb\right) dz \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(tb^2\right) \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(1-2t)\left(z^2 - 2z\frac{2tb}{(1-2t)}\right)\right) dz \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(tb^2\right) \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(1-2t)\left(z^2 - 2zq + q^2 - q^2\right)\right) dz \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(tb^2\right) \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(1-2t)\left((z-q)^2 - q^2\right)\right) dz \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(tb^2 + \frac{1}{2}(1-2t)\,q^2\right) \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(1-2t)\left((z-q)^2\right)\right) dz \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(\frac{tb^2}{1-2t}\right) \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left(\frac{z-q}{1/\sqrt{1-2t}}\right)^2\right) dz \\
&= \exp\left(\frac{tb^2}{1-2t}\right)(1-2t)^{-1/2}.
\end{aligned}
$$

Solution to Problem 7.19:

a) First notice that the density is symmetric. Make the transformation $y = x^{2n}$, $x = y^{\frac{1}{2n}}$, $dx = \frac{1}{2n} y^{\frac{1}{2n}-1} dy$ to get

$$
\int_0^{\infty} \frac{2dx}{(1+x^{2n})^m} = \int_0^{\infty} \frac{dy}{n\,(1+y)^m\, y^{1-\frac{1}{2n}}},
$$

followed by the substitution $z = (1+y)^{-1}$, $y = (1-z)\,z^{-1}$, $dy = -z^{-2}dz$, to get

$$
\begin{aligned}
-\frac{1}{n}\int_1^0 z^m z^{1-\frac{1}{2n}} (1-z)^{\frac{1}{2n}-1} z^{-2} dz &= \int_0^1 z^{\left(m-\frac{1}{2n}\right)-1} (1-z)^{\frac{1}{2n}-1} dz \\
&= \frac{1}{n} B\left(m - \frac{1}{2n}, \frac{1}{2n}\right).
\end{aligned}
$$

Thus, it must be the case that $n > 0$ and $m > (2n)^{-1}$.

b)

$$
\mu_r = \int_{-\infty}^{\infty} |x|^r f_X(x; n, m)\, dx = 2k \int_0^{\infty} \frac{x^r dx}{(1+x^{2n})^m},
$$

which, with the substitution

$$
z = \frac{1}{(1+x^{2n})} \qquad x = \left(\frac{1-z}{z}\right)^{\frac{1}{2n}} = \left(z^{-1}-1\right)^{\frac{1}{2n}} \qquad dx = -\frac{1}{2n}\left(\frac{1-z}{z}\right)^{\frac{1}{2n}-1} z^{-2}dz
$$

gives

$$
\begin{aligned}
\mu_r &= \frac{k}{n}\int_0^1 z^m \left(\frac{1-z}{z}\right)^{\frac{r}{2n}} \left(\frac{1-z}{z}\right)^{\frac{1}{2n}-1} z^{-2}dz \\
&= \frac{k}{n}\int_0^1 z^{\left(m-\frac{r}{2n}-\frac{1}{2n}\right)-1} (1-z)^{\frac{r}{2n}+\frac{1}{2n}-1} \\
&= \frac{k}{n} B\left(m - \frac{r}{2n} - \frac{1}{2n}, \frac{r}{2n} + \frac{1}{2n}\right) = \frac{\Gamma\left(m - \frac{r+1}{2n}\right)\Gamma\left(\frac{r+1}{2n}\right)}{\Gamma\left(m - \frac{1}{2n}\right)\Gamma\left(\frac{1}{2n}\right)},
\end{aligned}
$$

which exists for $2r < 2nm - 1$.

Solution to Problem 7.20: The density of X is given by

$$f_X(x;\, n_1, n_2) = \frac{1}{B\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} \left(\frac{n_1}{n_2}\right)^{n_1/2} \frac{x^{(n_1-2)/2}}{\left(1+\frac{n_1}{n_2}x\right)^{(n_1+n_2)/2}} \mathbb{I}_{(0,\infty)}(x).$$

Defining $a = n_1/n_2$, $\Pr(B \le b)$ is given by

$$\begin{aligned} \Pr\left(\frac{aX}{1+aX} \le b\right) &= \Pr\left(X \le \frac{b}{a(1-b)}\right) \\ &= \frac{1}{B(n_1/2, n_2/2)} a^{n_1/2} \int_0^{\frac{b}{a(1-b)}} \frac{x^{(n_1-2)/2}}{(1+ax)^{(n_1+n_2)/2}} dx. \end{aligned}$$

Then using the change of variable $y = y(x) = ax(1+ax)^{-1}$, then $x = ya^{-1}(1-y)^{-1}$ and $dx/dy = a^{-1}(1+y)^{-2}$ with the limits of integration $y(0) = 0$ and $y\left(\frac{b}{a(1-b)}\right) = b$ to get, noting that $1 + y(1-y)^{-1} = (1-y)^{-1}$,

$$\begin{aligned} \Pr(B \le b) &= \frac{1}{B(n_1/2, n_2/2)} a^{n_1/2} \int_0^b \frac{\left(\frac{y}{a(1-y)}\right)^{(n_1-2)/2}}{\left(1+\frac{ay}{a(1-y)}\right)^{(n_1+n_2)/2}} \frac{1}{a(1-y)^2} dy \\ &= \frac{1}{B(n_1/2, n_2/2)} a^{\left(\frac{n_1}{2} - \frac{n_1-2}{2} - 1\right)} \int_0^b y^{\frac{n_1}{2}-1} (1-y)^{-\frac{n_1-2}{2} + \frac{n_1+n_2}{2} - 2} dy \\ &= \frac{1}{B(n_1/2, n_2/2)} \int_0^b y^{\frac{n_1}{2}-1} (1-y)^{\frac{n_2}{2}-1} dy = \bar{B}_b(n_1/2, n_2/2), \end{aligned}$$

so that $B \sim \text{Beta}(n_1/2, n_2/2)$.

In general, if $X \sim \text{Beta}(a, b)$, then

$$f_X(x; a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} \mathbb{I}_{(0,1)}(x)$$

and, with $Y = 1 - X$,

$$f_Y(x; a, b) = f_X(1-y; a, b)\,|-1| = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} (1-y)^{a-1} y^{b-1} \mathbb{I}_{(0,1)}(1-y)$$

so that $Y \sim \text{Beta}(b, a)$. Thus,

$$1 - B = 1 - \frac{(n_1/n_2)X}{1+(n_1/n_2)X} = \frac{1}{1+(n_1/n_2)X} \sim \text{Beta}(n_2, n_1).$$

Solution to Problem 7.21: Calculating several derivatives of $f(a) = (1-a)^{-1/2}$ and using (7.36) gives

$$f^{(i)}(a) = \frac{1 \cdot 3 \cdot \cdots \cdot (2i-1)}{2^i} (1-a)^{-(i+1/2)} = \frac{\Gamma(i+1/2)}{\sqrt{\pi}} (1-a)^{-(i+1/2)},$$

so that the Taylor series expansion of f around zero is

$$\begin{aligned} (1-a)^{-1/2} &= \sum_{i=0}^{\infty} \frac{f^{(i)}(0)}{i!} a^i = \frac{1}{\sqrt{\pi}} \sum_{i=0}^{\infty} \frac{\Gamma(i+1/2)}{i!} a^i \\ &= 1 + \frac{1}{2}a + \frac{3}{8}a^2 + \frac{5}{16}a^3 + \frac{35}{128}a^4 + \frac{63}{256}a^5 + \cdots. \end{aligned}$$

Solution to Problem 7.22: Using (7.40), i.e., that

$$\lim_{x\to\infty} \bar{\Phi}(x) = \frac{1}{x\sqrt{2\pi}} \exp\left(-x^2/2\right),$$

we have

$$\begin{aligned}\lim_{x\to\infty}\frac{\bar{\Phi}\left(x+k/x\right)}{\bar{\Phi}\left(x\right)} &= \frac{\lim_{x\to\infty}\bar{\Phi}\left(x+k/x\right)}{\lim_{x\to\infty}\bar{\Phi}\left(x\right)}\\ &= \lim_{x\to\infty}\frac{x}{\left(x+k/x\right)}\exp\left(-\left(x+k/x\right)^2/2+x^2/2\right)\\ &= \lim_{x\to\infty}\frac{x}{\left(x+k/x\right)}\lim_{x\to\infty}\exp\left(-k\frac{2x^2+k}{2x^2}\right)\\ &= \exp\left(\lim_{x\to\infty}-k\frac{2x^2+k}{2x^2}\right)=\exp\left(-k\right).\end{aligned}$$

Solution to Problem 7.23: Try program S-7.2 with various values for `a`, `b` and `tol`, the numeric integration accuracy parameter, for example, `betacomp(1e-2,8,12)` and `betacomp(1e-7,8,12)`.

```
function betacomp(tol,a,b,yvec)
time1=[]; time2=[]; time3=[]; numint=[]; comb=[]; intern=[];
if nargin < 4, yvec=[0.05:0.05:0.95]; end
betafunc=beta(a,b); % just evaluate it once, then pass it as a parameter
for i=1:length(yvec)
  y=yvec(i);  uhr=cputime;
  numint=[numint quadl(@incbeta,0,y,tol,[],a,b,betafunc)];
  time1=[time1 cputime-uhr];

  uhr=cputime; j=a; n= b+j-1; kon=fact(n)/fact(n-j)/fact(j-1);
  comb=[comb (1-binocdf(j-1,n,y)) / kon / betafunc];
  time2=[time2 cputime-uhr];

  uhr=cputime; intern=[intern betainc(y,a,b)];
  time3=[time3 cputime-uhr];
end
subplot(3,1,1), plot(yvec,comb,'g-'); xlabel('y'), axis([0 1 0 1])
title(['Scaled IB function for a=',int2str(a),' and b=',int2str(b)]);

subplot(3,1,2), plot(yvec,time1,'r-',yvec,time2,'g--',yvec,time3,'b:');
xlabel('y'), legend('Int','Comb','Num',0); ylabel('seconds')
title(['Comp. of evaluation time of IB(',int2str(a),',',int2str(b),')'])

subplot(3,1,3), plot(yvec,comb-numint,'g-'); xlabel('y')
title ('Discrep. between combinatoric and numeric integration');

function f=fact(n), if n==1, f=1; else, f=n*fact(n-1); end

function f = incbeta(x,a,b,betafunc), f=x.^(a-1).*(1-x).^(b-1)/betafunc;
```

Program Listing S-7.2: Evaluation of the incomplete beta function

Solution to Problem 7.24: See Listings S-7.3 and S-7.4.

```
function cdfnorm(type,relative,x)
% x is vector of ordinates; type is any set of:
% 1:Polya (1945)   2:Hastings (1955) 3:Hart (1966)  4:Burr (1967)
% 5:Derenzo (1977) 6:Page (1977)     7:Moran (1980)
% if relative=1, computes relative error, otherwise absolute.

global true discrep;
if nargin<2, relative=1; end
if nargin<3, x=0.01:0.01:5; end
true=normcdf(x); lt=length(type);
if relative==1, etype='Relative'; else, etype='Absolute'; end
if lt==1
  [ameth,app,loc]=method(x,relative,type);
  subplot(2,1,1), plot(x,true,'g-',x,app,'r--'), grid
  title (['Exact F(x) (solid) and ',ameth,' Approximation G(x) (dashed)'])
  subplot(2,1,2), doplot(x,relative,loc,etype,ameth);
else
  switch lt
    case {2,3}, grow=lt; gcol=1;
    case 4,     grow=2; gcol=2;
    case {5,6}, grow=3; gcol=2;
    case {7,8}, grow=4; gcol=2;
    otherwise, error('Add more statements here'); end
  for i=1:length(type)
    [ameth,app,loc]=method(x,relative,type(i));
    subplot(grow,gcol,i), doplot(x,relative,loc,etype,ameth);
  end
end

function doplot(x,relative,loc,etype,ameth);
  global true discrep;
  plot(x,discrep,'r-')
  title (['     ',etype,' Discrepancy for ',ameth,' Approx.'])
  ax=axis; line([loc(2),loc(2)],[ax(3),ax(4)],'linestyle',':','color','b')
  line([ax(1),ax(2)],[0,0],'linestyle','-','color','k')
  xlabel(['Max = ',num2str(loc(1)),' at x = ',num2str(loc(2))])

function loc=getloc(app,x,relative);
  global true discrep;
  if relative==1, discrep=(true-app)./true; else, discrep=true-app; end
  ref=abs(discrep); worst=max(ref);
  bool=(worst==ref); worstx=x(bool); loc=[worst,worstx];
```

Program Listing S-7.3: Methods of approximating the normal cdf

```
function [ameth,app,loc]=method(x,relative,type);
  global discrep;
  switch type
    case 1
      ameth='Polya (1945)'; app=0.5*(1+sqrt(1-exp(-2*x.^2/pi)));
    case 2
      ameth='Hastings (1955)';
      t=(1+0.33267*x).^(-1); a1=0.4361836; a2=-0.1201676; a3=0.937298;
      app=1-normpdf(x).*(a1*t + a2*t.^2 + a3*t.^3);
    case 3
      ameth='Hart (1966)';
      x2=x.^2; n=1+sqrt(1+6*pi-2*pi^2);  a=n / (2*pi);   b=n^2 / (2*pi);
      c=x*sqrt(pi/2); d=1./(x*sqrt(2*pi));
      e=exp(-x2 / 2); f=sqrt(1+b*x2) ./ (1+a*x2);
      curly=(c+sqrt(c.^2 +f.*e)).^(-1); app=1-d.*e.*( 1-f.*curly );
    case 4
      ameth='Burr (1967)';
      a=0.644693; b=0.161984; c=4.874; k=-6.158;
      gx=1-(1+(a+b*x).^c).^k; app=gx; gmx=1-(1+(a+b*(-x)).^c).^k;
      app=real(0.5*(gx+1-gmx));
    case 5
      ameth='Derenzo (1977)';
      xlo=x(x<=5.5); xhi=x(x>5.5); num=(83*xlo+351).*xlo + 562;
      den=703./xlo + 165;  applo=1-0.5*exp(-num./den);
      apphi=1-1./sqrt(2*pi)./xhi .* exp(-0.5*xhi.^2 - 0.94./xhi.^2);
      app=[applo;apphi];
    case 6
      ameth='Page (1977)';
      a1=0.7988; a2=0.04417;
      y=a1*x.*(1+a2*x.^2); e=exp(2*y); app=e./(1+e);
    case 7
      ameth='Moran (1980)';
      s=0; h=sqrt(2)/3;
      for n=0:12
        t=n+0.5; k=exp(-t^2 / 9) / t; s=s+k*sin(h*t*x);
      end
      app=0.5+s/pi;
    otherwise, error('Unknown method.'), end
    loc=getloc(app,x,relative);
```

Program Listing S-7.4: Listing S-7.3 continued

Solutions to Chapter 8: Joint and Conditional Random Variables

Solution to Problem 8.1: Let ϕ_ν and Φ_ν denote the pdf and cdf of the Student's t distribution with ν degrees of freedom, respectively. With $K = v^{-1/2} \,/\, B(\nu/2, 1/2)$ and $u = 1 + r^2/\nu$,

$$\begin{aligned}
\mathbb{E}[R \mid R < c] &= \frac{1}{\Phi_\nu(c)} \int_{-\infty}^{c} r\phi_\nu(r)dr \\
&= \frac{K}{\Phi_\nu(c)} \int_{-\infty}^{c} r\left(1 + r^2/\nu\right)^{-(\nu+1)/2} dr = \frac{K}{\Phi_\nu(c)} \frac{\nu}{2} \int_{\infty}^{1+c^2/\nu} u^{-(\nu+1)/2} du \\
&= -\frac{K}{\Phi_\nu(c)} \frac{\nu}{1-\nu}\, u^{1-(\nu+1)/2}\Big|_{1+c^2/\nu}^{\infty} = \frac{K}{\Phi_\nu(c)} \frac{\nu}{1-\nu} \left(1 + c^2/\nu\right)^{1-(\nu+1)/2} \\
&= -\frac{\phi_\nu(c)}{\Phi_\nu(c)} \times \left[\frac{\nu + c^2}{\nu - 1}\right],
\end{aligned}$$

from which it is clear that, as $\nu \to \infty$, the expression approaches that based on the normal distribution given in Example 8.15.

Solution to Problem 8.2:

a) The event of interest is

$$\begin{aligned}
E &= (A_1 \text{ has } \nu_1 \text{ hearts } \cap\ A_2 \text{ has } \nu_2 \text{ hearts}) \\
&\quad \cup (A_1 \text{ has } \nu_2 \text{ hearts } \cap\ A_2 \text{ has } \nu_1 \text{ hearts}).
\end{aligned}$$

There are $\binom{13}{\nu_1}\binom{52-13}{13-\nu_1}$ ways for A_1 to get ν_1 hearts, leaving $\binom{13-\nu_1}{\nu_2}\binom{52-13-(13-\nu_1)}{13-\nu_2}$ ways for A_2 to get ν_2 hearts. We are not interested in the number of hearts that B_1 and B_2 possess, so that there are simply $\binom{26}{13,13}$ ways of distributing the remaining 26 cards among B_1 and B_2. There are $\binom{52}{13,13,13,13}$ ways in total of distributing the cards, but $\binom{52}{13,13,13,13} = \binom{52}{13}\binom{39}{13}\binom{26}{13}$, and $\binom{26}{13,13} = \binom{26}{13}$ so that defining k to be

$$k_{\nu_1,\nu_2} = \frac{\binom{13}{\nu_1}\binom{52-13}{13-\nu_1}\binom{13-\nu_1}{\nu_2}\binom{52-13-(13-\nu_1)}{13-\nu_2}\binom{26}{13,13}}{\binom{52}{13,13,13,13}} = \frac{\binom{13}{\nu_1}\binom{39}{13-\nu_1}\binom{13-\nu_1}{\nu_2}\binom{26+\nu_1}{13-\nu_2}}{\binom{52}{13}\binom{39}{13}}$$

we see

$$\Pr(E) = \begin{cases} 2k_{\nu_1,\nu_2} & \text{if } \nu_1 \neq \nu_2 \text{ and } \nu_1 + \nu_2 \leq 13 \\ k_{\nu_1,\nu_1} & \text{if } \nu_1 = \nu_2 \text{ and } \nu_1 + \nu_2 \leq 13 \\ 0 & \text{otherwise.} \end{cases}$$

b) From the previous problem, this must be given by $S_\nu = \sum_{\nu_1=0}^{\nu} k_{\nu_1,\nu-\nu_1}$ for ν even or odd. Instead of trying to simplify S_ν, we can proceed directly as follows. There are $\binom{13}{\nu}$ ways of picking the ν hearts and, thus, $\binom{39}{26-\nu}$ ways of picking the remaining cards for A_1 and A_2.

Then, there are $\binom{26}{13,13}$ ways of distributing the 26 chosen cards among A_1 and A_2. From the remaining cards, there are $\binom{26}{13,13}$ ways of distributing them among B_1 and B_2, so that

$$S_\nu = \frac{\binom{13}{\nu}\binom{39}{26-\nu}\binom{26}{13,13}\binom{26}{13,13}}{\binom{52}{13,13,13,13}} = \frac{\binom{13}{\nu}\binom{39}{26-\nu}\binom{26}{13}}{\binom{52}{13}\binom{39}{13}}.$$

c) $\Pr(C \mid D) = \Pr(CD) / \Pr(D) = \Pr(C) / \Pr(D)$, but event C is exactly the question in part (**a**), and D is exactly the question in part (**b**), so the answer is

$$\Pr(C \mid D) = S_\nu^{-1} \begin{cases} 2k_{\nu_1,\nu_2}, & \text{if } \nu_1 \neq \nu_2 \text{ and } \nu_1 + \nu_2 \leq 13, \\ k_{\nu_1,\nu_2}, & \text{if } \nu_1 = \nu_2 \text{ and } \nu_1 + \nu_2 \leq 13, \\ 0, & \text{otherwise.} \end{cases}$$

Another solution is to work directly with the reduced sample space, i.e., given that between the two people of interest, they have ν hearts. Observe that the probability that A_1 gets ν_1 hearts is hypergeometrically distributed, i.e.,

$$\Pr(A_1 \text{ gets } \nu_1 \text{ hearts}) = \frac{\binom{\nu}{\nu_1}\binom{26-\nu}{13-\nu_1}}{\binom{26}{13}}.$$

Thus, we also have

$$\Pr(C \mid D) = \binom{26}{13}^{-1} \begin{cases} 2\binom{\nu}{\nu_1}\binom{26-\nu}{13-\nu_1}, & \text{if } \nu_1 \neq \nu_2 \text{ and } \nu_1 + \nu_2 \leq 13, \\ \binom{\nu}{\nu_1}\binom{26-\nu}{13-\nu_1}, & \text{if } \nu_1 = \nu_2 \text{ and } \nu_1 + \nu_2 \leq 13, \\ 0, & \text{otherwise,} \end{cases}$$

which is easier to evaluate numerically. For $\nu_1 = 3$ and $\nu_2 = 2$, $\Pr(C \mid D) = \frac{78}{115} \approx 0.678$.

d) Intuitively, on average the ν hearts should be equally distributed between A_1 and A_2, so that $\mathbb{E}[\nu_1 \mid \nu] = \frac{\nu}{2}$. Formally,

$$\mathbb{E}[\nu_1 \mid \nu] = \frac{\sum_{\nu_1=0}^{\nu} \nu_1 \cdot k_{\nu_1,\nu-\nu_1}}{\sum_{i=0}^{\nu} k_{i,\nu-i}} = \frac{\sum_{\nu_1=0}^{\nu} \nu_1 \binom{13}{\nu_1}\binom{39}{13-\nu_1}\binom{13-\nu_1}{\nu-\nu_1}\binom{26+\nu_1}{13-(\nu-\nu_1)}}{\binom{13}{\nu}\binom{39}{26-\nu}\binom{26}{13}} = \frac{\nu}{2},$$

where the last equality can be verified numerically for $1 \leq \nu \leq 13$ or, possibly, proven by induction. Use of the hypergeometric distribution in part (**c**) simplifies things considerably, so that, directly,

$$\mathbb{E}[\nu_1 \mid \nu] = \frac{13\nu}{26} = \frac{\nu}{2}, \quad \mathbb{V}(\nu_1 \mid \nu) = \frac{13\nu}{26} \cdot \frac{26-\nu}{26} \cdot \frac{26-13}{26-1} = \frac{\nu(26-\nu)}{100}.$$

Solution to Problem 8.3: From Example 8.17, $\Pr(X_1 < aX_2) = \frac{a\lambda_1}{a\lambda_1+\lambda_2}$, so that

$$\Pr(X_1 < aX_2 \mid X_1 < bX_2) = \frac{\Pr(X_1 < aX_2)}{\Pr(X_1 < bX_2)} = \frac{a}{b}\frac{b\lambda_1 + \lambda_2}{a\lambda_1 + \lambda_2}.$$

As b gets large, irrespective of the λ_i, it will eventually be "almost certain" that $X_1 < bX_2$, so that the conditional probability reduces to just $\Pr(X_1 < aX_2)$. Writing the above conditional probability as

$$\frac{a\lambda_1 + a\lambda_2/b}{a\lambda_1 + \lambda_2}$$

and taking the limit as $b \to \infty$ immediately confirms this.

Solution to Problem 8.4: The density is just

$$f_{X,Y}(x,y) = \mathbb{I}_{(0,1)}(x)\,\mathbb{I}_{(0,1)}(y).$$

Note that the "solution"

$$\Pr(Z \le z) = \Pr\left(Y \le \frac{z}{X}\right) = \int_0^1 \int_0^{z/x} dydx = z\int_0^1 \frac{dx}{x} = \infty$$

is wrong. The reason is that, for $X \le z$, $\Pr(Y \le z/X) = 1$. By splitting up the region into two pieces, we obtain

$$F_Z(z) = \int_0^z \int_0^1 dydx + \int_z^1 \int_0^{z/x} dydx = (z - z\ln z)\,\mathbb{I}_{(0,1)}(z)$$

and differentiating, $f_Z(z) = -\ln(z)\,\mathbb{I}_{(0,1)}(z)$.

Solution to Problem 8.5:

a) From the discussion in §4.5 relating Poisson and exponential r.v.s, $\Pr\left(S_1^1 < S_1^2\right)$ is the same as $\Pr(X_1 < X_2)$ for $X_i \stackrel{\text{ind}}{\sim} \operatorname{Exp}(\lambda_i)$, $i = 1, 2$. From Example 8.17, this is given by $\frac{\lambda_1}{\lambda_1+\lambda_2}$.

b) This means that the first event must be from N_1, which occurs with probability $\frac{\lambda_1}{\lambda_1+\lambda_2}$ and that the second event also must be from N_1, which also occurs with probability $\frac{\lambda_1}{\lambda_1+\lambda_2}$ from the memoryless property of the exponential, so that

$$\Pr\left(S_n^1 < S_1^2\right) = \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^n.$$

c) Similar to the previous part, the next event occurs from process N_1 with probability $\frac{\lambda_1}{\lambda_1+\lambda_2}$, and from N_2 with probability $1 - \frac{\lambda_1}{\lambda_1+\lambda_2} = \frac{\lambda_2}{\lambda_1+\lambda_2}$.

d) At any given time, when the next event occurs, it is from N_i with probability $\frac{\lambda_i}{\lambda_1+\lambda_2}$, $i = 1, 2$. Thus, if we think of the 2 processes as "competing" in the sense that, at each occurrence, N_i "wins" with probability $\frac{\lambda_i}{\lambda_1+\lambda_2}$, $i = 1, 2$, we can use (3.18) so that, with $p = \frac{\lambda_1}{\lambda_1+\lambda_2}$,

$$\Pr\left(S_n^1 < S_m^2\right) = \sum_{k=n}^{n+m-1} \binom{n+m-1}{k} p^k (1-p)^{n+m-1-k}.$$

Solution to Problem 8.6:

a) With $Y = y$ known and assuming that we draw without replacement (as is appropriate for a lottery), X follows the hypergeometric distribution

$$f_{X|Y}(x \mid y) = \frac{\binom{y}{x}\binom{N-y}{n-x}}{\binom{N}{n}} \mathbb{I}_{\{0,1,\ldots,\min(n,y)\}}(x)$$

with

$$\mathbb{E}[X \mid Y = y] = \frac{ny}{N}. \tag{S-8.1}$$

b) From the definition,

$$\begin{aligned} f_{X,Y}(x,y) &= f_Y(y)\, f_{X|Y}(x \mid y) \\ &= \binom{N}{y}\theta^y (1-\theta)^{N-y} \frac{\binom{y}{x}\binom{N-y}{n-x}}{\binom{N}{n}} \mathbb{I}_{\{0,1,\ldots,n\}}(y)\,\mathbb{I}_{\{0,1,\ldots,\min(n,y)\}}(x) \end{aligned}$$

and, simplifying the combinatorics and noting the explicit dependence on n and N,

$$f_{X,Y}(x, y \mid n, N) = \binom{n}{x}\binom{N-n}{y-x}\theta^y (1-\theta)^{N-y}\, \mathbb{I}_{\{0,1,\ldots,n\}}(y)\,\mathbb{I}_{\{0,1,\ldots,\min(n,y)\}}(x).$$

c) Multiplying the left hand side by $1 = \theta^x \theta^{-x} (1-\theta)^{n-x} (1-\theta)^{-(n-x)}$ yields

$$\theta^x (1-\theta)^{n-x} \sum_{y=0}^{N} \binom{N-n}{y-x} \theta^{y-x} (1-\theta)^{N-n-(y-x)},$$

with the sum only being valid from $y = x$ to $y = N - n + x$ so that the combinatoric is defined. (This makes sense, observing that $y < x$ would imply having drawn more winning tickets than issued, and $N - n < y - x \Leftrightarrow N - y < n - x$ would imply having drawn more losing tickets than issued). Restricting the sum to this range, and defining $z = y - x$, we have

$$\sum_{y=x}^{N-n+x} \binom{N-n}{y-x} \theta^{y-x} (1-\theta)^{N-n-(y-x)} = \sum_{z=0}^{N-n} \binom{N-n}{z} \theta^{z} (1-\theta)^{(N-n)-z},$$

with the right hand side being the sum of all the binomial coefficients, which is one, as was to be shown.

d) $f_{Y|X}(y \mid x) = f_{X,Y}(x,y) \,/\, f_X(x)$, where

$$\begin{aligned} f_X(x) &= \sum_{y=0}^{N} f_{X,Y}(x,y) \\ &= \sum_{y=0}^{N} \binom{n}{x} \binom{N-n}{y-x} \theta^y (1-\theta)^{N-y} \mathbb{I}_{\{0,1,\ldots,n\}}(y) \, \mathbb{I}_{\{0,1,\ldots,\min(n,y)\}}(x) \\ &= \binom{n}{x} \theta^x (1-\theta)^{n-x} \mathbb{I}_{\{0,1,\ldots,n\}}(x) \qquad \text{(S-8.2)} \end{aligned}$$

and

$$f_{Y|X}(y \mid x) = \binom{N-n}{y-x} \theta^{y-x} (1-\theta)^{N-n-(y-x)} \mathbb{I}_{\{x,\ldots,N-n+x\}}(y).$$

e) The density of $f_{Y|X}$ is a "shifted" binomial, i.e., it starts at x instead of 0. We might have guessed at the answer by assuming that the binomial form still holds, but using the obvious facts that $y \geq x$ and $N - y \geq n - x$.

f) Defining $z = y - x$ and using the mean of the binomial distribution,

$$\begin{aligned} \mathbb{E}[Y \mid X] &= \sum_{y=x}^{N-n+x} y \binom{N-n}{y-x} \theta^{y-x} (1-\theta)^{N-n-(y-x)} \\ &= \sum_{z=0}^{N-n} (z+x) \binom{N-n}{z} \theta^z (1-\theta)^{N-n-z} \\ &= x + (N-n)\,\theta. \qquad \text{(S-8.3)} \end{aligned}$$

For $n = N$ we observe the whole sample and x, the number of observed successes, must be y. This agrees with $\mathbb{E}[Y \mid X, N = n] = x$. For $n = 0$, x is naturally also zero, so that $\mathbb{E}[Y \mid X] = N\theta = \mathbb{E}[Y]$, the *unconditional* expectation of Y.

g)

i) X is binomially distributed and, hence, $\mathbb{E}[X] = n\theta$.

ii) As Y is unconditionally binomially distributed,

$$N\theta = \mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y \mid X]] = \mathbb{E}[x] + (N-n)\,\theta$$

so that

$$\mathbb{E}[X] = N\theta - (N-n)\,\theta = n\theta.$$

iii) $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid Y]] = \mathbb{E}[nY/N] = nN\theta/N = n\theta$.

Solution to Problem 8.7:

a) As floors are equally likely to be chosen,

$$\mathbb{E}\left[I_1 \mid X = x\right] = \cdots = \mathbb{E}\left[I_N \mid X = x\right],$$

so that taking iterative expectations yields

$$\begin{aligned}\mathbb{E}\left[Y\right] &= \mathbb{E}_X\left[\mathbb{E}_{Y|X}\left[Y \mid X = x\right]\right] \\ &= \mathbb{E}_X\left[\sum_{i=1}^{N}\mathbb{E}\left[I_i \mid X = x\right]\right] = N\mathbb{E}_X\left[\mathbb{E}\left[I_1 \mid X = x\right]\right].\end{aligned}$$

As

$$\begin{aligned}\mathbb{E}\left[I_1 \mid X = x\right] &= \Pr\left(\text{at least one person gets off on } 1^{\text{st}} \text{ floor}\right) \\ &= 1 - \Pr\left(\text{no one gets off on } 1^{\text{st}} \text{ floor}\right) \\ &= 1 - \left(\frac{N-1}{N}\right)^x,\end{aligned}$$

it follows that

$$\begin{aligned}\mathbb{E}\left[Y\right] = N\mathbb{E}_X\left[1 - \left(\frac{N-1}{N}\right)^x\right] &= N\sum_{x=0}^{\infty}\left(1 - \left(\frac{N-1}{N}\right)^x\right)\frac{e^{-\lambda}\lambda^x}{x!} \\ &= N\left(1 - e^{-\lambda}\sum_{x=0}^{\infty}\left(\frac{N-1}{N}\right)^x\frac{\lambda^x}{x!}\right).\end{aligned}$$

But directly from the Poisson distribution,

$$\sum_{x=0}^{\infty}\frac{\left(\lambda\left(\frac{N-1}{N}\right)\right)^x}{x!} = 1/e^{-\lambda\left(\frac{N-1}{N}\right)} = e^{\lambda}e^{-\lambda/N}$$

so that $\mathbb{E}\left[Y\right] = N\left(1 - e^{-\lambda/N}\right)$.

b) With only one floor, $\mathbb{E}\left[Y\right]$ reduces to the weighted probability $0\times\Pr\left(X = 0\right) + 1\times\Pr\left(X \geq 1\right) = 1 - \Pr\left(X = 0\right) = 1 - e^{-\lambda}\lambda^0/0! = 1 - e^{-\lambda}$ which approaches one as λ increases, i.e., $\Pr\left(X = 0\right) \to 0$.

c) Because people choose floors independently and each floor is equally likely, as the number of floors gets large, the probability that two or more people pick the same floor will get smaller, so that $\lim_{N\to\infty}\mathbb{E}\left[Y\right] = \mathbb{E}\left[X\right] = \lambda$.

d) The series expansion of $N\left(1 - e^{-\lambda/N}\right)$ is

$$N\left(1 - \left(1 - \frac{\lambda}{N} + \frac{\lambda^2}{2N^2} - \frac{\lambda^3}{3!N^3} + \cdots\right)\right) = \lambda - \frac{\lambda^2}{2N} + \frac{\lambda^3}{3!N^2} - \cdots$$

which is equal to λ for large N. Alternatively, from L´Hospitals rule with the substitution $M = 1/N$ gives

$$\lim_{M\to 0}\frac{\left(1 - e^{-\lambda M}\right)}{M} = \lim_{M\to 0}\frac{\lambda e^{-\lambda M}}{1} = \lambda.$$

Solution to Problem 8.8: For density $2y\left(2 - x^2 - y\right)\mathbb{I}_{(0,1)}\left(x\right)\mathbb{I}_{(0,1)}\left(y\right)$ as shown in Figure S-8.1,

$$\begin{aligned}f_X\left(x\right) &= \int_{-\infty}^{\infty} f_{X,Y}\left(x,y\right)dy = 2\int_0^1 y\left(2 - x^2 - y\right)dy = \left(\frac{4}{3} - x^2\right)\mathbb{I}_{(0,1)}\left(x\right), \\ f_Y\left(y\right) &= \int_{-\infty}^{\infty} f_{X,Y}\left(x,y\right)dx = 2\int_0^1 y\left(2 - x^2 - y\right)dx = \frac{2}{3}y\left(5 - 3y\right)\mathbb{I}_{(0,1)}\left(y\right)\end{aligned}$$

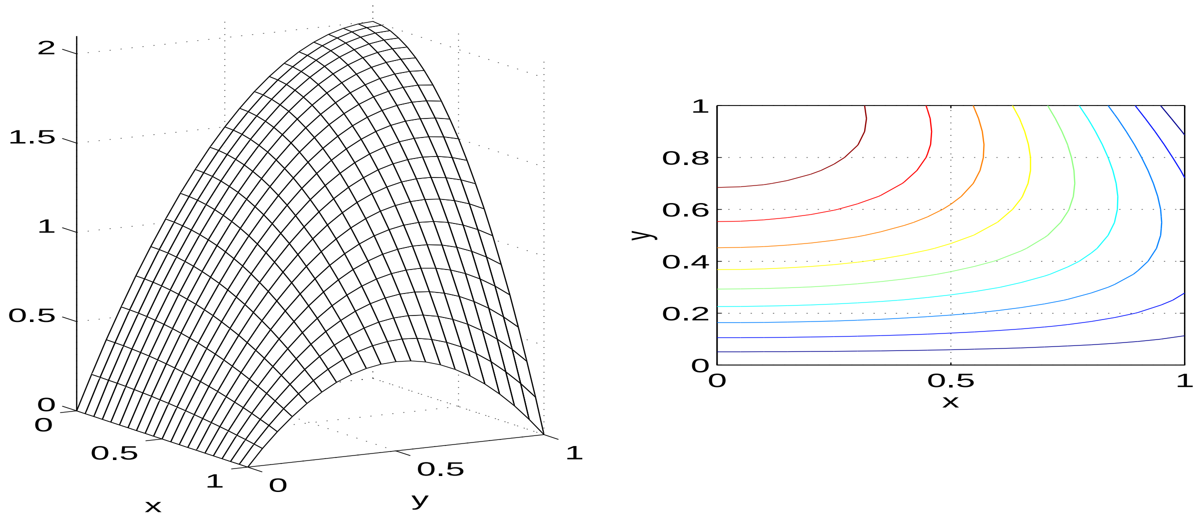


Figure S-8.1: Density $2y\left(2-x^2-y\right)\mathbb{I}_{(0,1)}\left(x\right)\mathbb{I}_{(0,1)}\left(y\right)$

and

$$f_{X|Y}\left(x\mid y\right)=3\frac{2-x^2-y}{5-3y}\mathbb{I}_{(0,1)}\left(x\right)\mathbb{I}_{(0,1)}\left(y\right),$$

$$f_{Y|X}\left(y\mid x\right)=6y\frac{2-x^2-y}{4-3x^2}\mathbb{I}_{(0,1)}\left(y\right)\mathbb{I}_{(0,1)}\left(x\right).$$

These are shown in Figures S-8.2 and S-8.3.

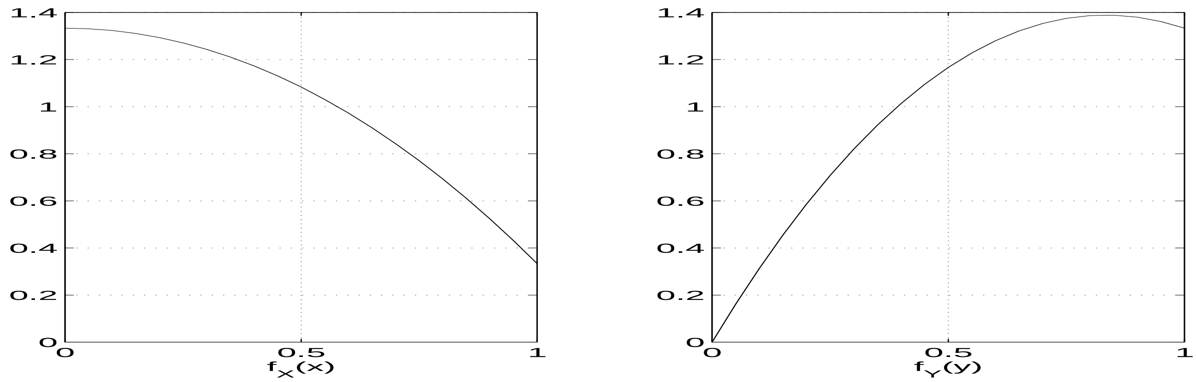


Figure S-8.2: Marginal densities

Also,

$$\begin{aligned}f_X\left(x\right) &= \int_{-\infty}^{\infty}f_{X|Y}\left(x\mid y\right)f_Y\left(y\right)dy\\ &= \int_0^1 3\frac{2-x^2-y}{5-3y}\frac{2}{3}y\left(5-3y\right)dy=\left(\frac{4}{3}-x^2\right)\mathbb{I}_{(0,1)}\left(x\right)\end{aligned}$$

and

$$\begin{aligned}f_{X|Y}\left(x\mid y\right) &= \frac{f_X\left(x\right)f_{Y|X}\left(y\mid x\right)}{\int_{-\infty}^{\infty}f_{Y|X}\left(y\mid x\right)dF_X\left(x\right)}=\frac{\left(\frac{4}{3}-x^2\right)6y\frac{2-x^2-y}{4-3x^2}}{\int_0^1\left(\frac{4}{3}-x^2\right)6y\frac{2-x^2-y}{4-3x^2}dx}\\ &= 3\frac{2-x^2-y}{5-3y}\mathbb{I}_{(0,1)}\left(x\right)\mathbb{I}_{(0,1)}\left(y\right).\end{aligned}$$

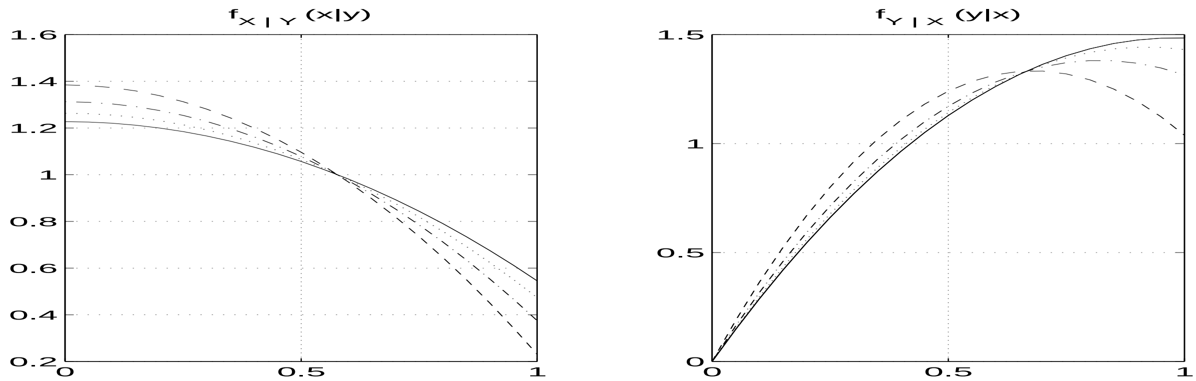


Figure S-8.3: Conditional densities; on left, $y = 0.2$ (solid), $y = 0.4$ (dotted), $y = 0.6$ (dashdot), $y = 0.8$ (dashed); on right $x = 0.2$ (solid), $x = 0.4$ (dotted), $x = 0.6$ (dashdot), $x = 0.8$ (dashed)

Solution to Problem 8.9: With $u = x/y$,

$$f_Y(y) = ce^{-y}\int_0^1 (uy)^{a-1}(y(1-u))^{b-1}\,y\,du = ce^{-y}y^{a+b-1}B(a,b)\,\mathbb{I}_{(0,\infty)}(y).$$

From the kernel $e^{-y}y^{a+b-1}$, we see that $Y \sim \text{Gam}(a+b,1)$ and c is derived from

$$1 = \int_0^\infty f_Y(y)\,dy = c\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}\int_0^\infty e^{-y}y^{a+b-1}dy = c\Gamma(a)\Gamma(b).$$

For X, with $u = y - x$, $y = u + x$,

$$\begin{aligned} f_X(x) &= \frac{x^{a-1}}{\Gamma(a)\Gamma(b)}\int_x^\infty (y-x)^{b-1}e^{-y}dy = \frac{x^{a-1}e^{-x}}{\Gamma(a)\Gamma(b)}\int_0^\infty u^{b-1}e^{-u}du \\ &= \frac{1}{\Gamma(a)}x^{a-1}e^{-x}\mathbb{I}_{(0,\infty)}(x), \end{aligned}$$

i.e., $X \sim \text{Gam}(a,1)$.

The conditionals simplify to

$$\begin{aligned} f_{X|Y}(x \mid y) &= \frac{1}{B(a,b)}x^{a-1}(y-x)^{b-1}y^{1-a-b}\mathbb{I}_{(0,y)}(x), \\ f_{Y|X}(y \mid x) &= \frac{1}{\Gamma(b)}(y-x)^{b-1}e^{x-y}\mathbb{I}_{(x,\infty)}(y). \end{aligned}$$

Note the special cases $(X \mid Y = 1) \sim \text{Beta}(a,b)$ and $(Y \mid X = 0) \sim \text{Gam}(b,1)$.

Solution to Problem 8.10: Observe that $(X \mid N = n) \sim \text{Bin}(n,p)$ so that, from (8.16) with $q = 1-p$,

$$\begin{aligned} \Pr(X = x) &= \sum_{n=x}^\infty \Pr(X = x \mid N = n)\Pr(N = n) \\ &= \sum_{n=x}^\infty \binom{n}{x}p^x(1-p)^{n-x}\frac{e^{-\lambda}\lambda^n}{n!} \\ &= \frac{p^xe^{-\lambda}}{x!}\sum_{n=x}^\infty \frac{\lambda^n(1-p)^{n-x}}{(n-x)!} \\ &= \frac{p^xe^{-\lambda}}{x!}\left(\lambda^x + \lambda^{x+1}q + \frac{1}{2}\lambda^{x+2}q^2 + \cdots\right) \\ &= \frac{p^xe^{-\lambda}}{x!}\lambda^x\left(1 + \lambda q + \frac{1}{2}\lambda^2q^2 + \cdots\right) = \frac{p^xe^{-\lambda}}{x!}\lambda^xe^{\lambda q} \\ &= \frac{(\lambda p)^x e^{-(\lambda p)}}{x!}\mathbb{I}_{\{0,1,\ldots\}}(x). \end{aligned}$$

Thus, unconditionally, $X \sim \text{Poi}(\lambda p)$. This also confirms $\mathbb{E}[X]$ and $\mathbb{V}(X)$ computed in Example 8.14.

Solution to Problem 8.11:

a) $f_Y(y) = e^{-y} \int_0^\infty e^{-x/y} y^{-1} dx = e^{-y} \mathbb{I}_{(0,\infty)}(y)$.

b) $f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = e^{-x/y} y^{-1} \mathbb{I}_{(0,\infty)}(x)$ for $y > 0$.

c) $\Pr(X > Y) = \int_0^\infty \int_y^\infty f_{X,Y}(x,y)\, dx dy = \int_0^\infty \int_0^x f_{X,Y}(x,y)\, dy dx$ and, using the former,

$$\int_0^\infty \int_y^\infty e^{-x/y} e^{-y} y^{-1} dx dy = \int_0^\infty e^{-y} \left[\int_y^\infty e^{-x/y} y^{-1} dx \right] dy.$$

With $a = 1/y$, the inner integral is $\int_y^\infty a e^{-ax} dx = 1 - F(y) = e^{-ay} = e^{-1}$, so that

$$\Pr(X > Y) = \int_0^\infty e^{-y} e^{-1} dy = e^{-1}.$$

Solution to Problem 8.12: Based on density (8.21) and covariance expression (5.24),

$$\begin{aligned}
\text{Cov}(X,Y) &= \int_0^\infty \int_0^\infty xy f(x,y;\theta)\, dx\, dy - \mathbb{E}[X]\mathbb{E}[Y] \\
&= \int_0^\infty \int_0^\infty xy\left((1+\theta x)(1+\theta y) - \theta\right) \exp(-x - y - \theta xy)\, dx\, dy - 1 \\
&= \int_0^\infty y e^{-y} \left[\int_0^\infty x\left((1+\theta x)(1+\theta y) - \theta\right) e^{-x(1+\theta y)} dx \right] dy - 1 \\
&= \int_0^\infty y e^{-y} A(y)\, dy - 1,
\end{aligned}$$

where $A(y)$ is the inner integral, given by

$$\begin{aligned}
A(y) &= \int_0^\infty x\left((1+\theta x)(1+\theta y) - \theta\right) e^{-x(y\theta+1)}\, dx \\
&= (1+\theta y)\left[\int_0^\infty x e^{-x(1+\theta y)}\, dx + \theta \int_0^\infty x^2 e^{-x(1+\theta y)}\, dx \right] - \theta \int_0^\infty x e^{-x(1+\theta y)}\, dx \\
&= (1+\theta y)\left(\frac{1}{(1+\theta y)^2} + \frac{2\theta}{(1+\theta y)^3} \right) - \frac{\theta}{(1+\theta y)^2} \\
&= \frac{1}{1+\theta y} + \frac{\theta}{(1+\theta y)^2}
\end{aligned}$$

having used the elementary calculations

$$\int_0^\infty x e^{-x(1+y\theta)}\, dx = (1+\theta y)^{-2} \quad \text{and} \quad \int_0^\infty x^2 e^{-x(1+\theta y)}\, dx = 2(1+\theta y)^{-3}.$$

Solution to Problem 8.13:

a) With $G_m := \mathbb{E}\left[N_m^2\right] = \mathbb{E}\left[\mathbb{E}\left[N_m^2 \mid N_{m-1}\right]\right]$ from the law of the iterated expectation, define $X := N_m^2 \mid (N_{m-1} = n)$ and let Y the $(n+1)^{\text{st}}$ trial. Then, similar to the calculation and argumentation in Example 8.13,

$$\begin{aligned}
\mathbb{E}[X] &= (1-p) \cdot \mathbb{E}[X \mid Y = 0] + p \cdot \mathbb{E}[X \mid Y = 1] \qquad \text{(S-8.4)} \\
&= (1-p) \cdot \mathbb{E}\left[(N_m + n + 1)^2\right] + p \cdot (n+1)^2.
\end{aligned}$$

Observe that n is a constant (and not a random variable), so that expanding the square, we can write

$$\begin{aligned}\mathbb{E}[X] &= (1-p)\,\mathbb{E}\left[N_m^2\right] + 2(1-p)\,n\mathbb{E}[N_m] + 2(1-p)\,\mathbb{E}[N_m] \\ &\quad + (1-p)(n+1)^2 + p\cdot(n+1)^2 \\ &= (1-p)\,G_m + 2(1-p)(n+1)\,\mathbb{E}[N_m] + (n+1)^2 .\end{aligned}$$

Now taking expectations of both sides w.r.t. N_{m-1} (i.e., n) and noting that G_m and $\mathbb{E}[N_m]$ are constants gives

$$G_m = (1-p)\,G_m + 2(1-p)\,\mathbb{E}[N_m]\,\mathbb{E}[(N_{m-1}+1)] + \mathbb{E}\left[(N_{m-1}+1)^2\right]$$

or, simplifying and letting $E_m := \mathbb{E}[N_m]$,

$$p\,G_m = 2(1-p)\,E_m(1+E_{m-1}) + G_{m-1} + 2E_{m-1} + 1.$$

With $m=1$, $N_1 \sim \text{Geo}(p)$ and, from (4.51),

$$G_1 = \mathbb{E}\left[N_1^2\right] = \frac{1-p}{p^2} + \left(\frac{1}{p}\right)^2 = \frac{2-p}{p^2}.$$

Then, using the recursion and simplifying, we get

$$\begin{aligned}G_1 &= p^{-2}(-p+2), \\ G_2 &= p^{-4}\left(-p^3-p^2+4p+2\right), \\ G_3 &= p^{-6}\left(-p^5-p^4-p^3+6p^2+4p+2\right), \\ G_4 &= p^{-8}\left(-p^7-p^6-p^5-p^4+8p^3+6p^2+4p+2\right)\end{aligned}$$

for which a general pattern seems to be given by

$$\begin{aligned}G_m &= p^{-2m}\left(-\sum_{i=m}^{2m-1} p^i + 2\sum_{i=0}^{m-1}(i+1)\,p^i\right) \\ &= \frac{p^{2m} - (3+2m)\,p^m - p^{2m+1} + (1+2m)\,p^{m+1} + 2}{(1-p)^2\,p^{2m}},\end{aligned}$$

which is (8.51). The expression $\mathbb{V}(N_m) = G_m - E_m^2$ then reduces to (8.49), i.e.,

$$\mathbb{V}(N_m) = \frac{1-(1+2m)(1-p)\,p^m - p^{2m+1}}{(1-p)^2\,p^{2m}}.$$

Substituting the identity $E_{m-1} = pE_m - 1$ from (8.32) into (8.50) gives

$$G_m = \frac{1}{p}G_{m-1} + 2E_m - \frac{1}{p} + 2E_m^2(1-p) = \frac{1}{p}G_{m-1} + \frac{2}{1-p}\left(p^{-2m} - p^{-m}\right) - \frac{1}{p}.$$

As G_0 is zero, recursively substituting and simplifying yields

$$\begin{aligned}G_m &= \sum_{i=1}^{m}\left(\frac{1}{p}\right)^{m-i}\left(\frac{2}{1-p}\right)\left(p^{-2i} - p^{-i}\right) - \sum_{i=1}^{m}\left(\frac{1}{p}\right)^{i} \\ &= \left(\frac{2}{1-p}\right)\left(\frac{p^{-2m}-p^{-m}}{1-p} - mp^{-m}\right) - \frac{p^{-m}-1}{1-p},\end{aligned}$$

which the reader can verify is equivalent to (8.51).

b) With $M_m := \mathbb{E}\left[e^{tN_m}\right] = \mathbb{E}\left[\mathbb{E}\left[e^{tN_m} \mid N_{m-1}\right]\right]$ and $X = e^{tN_m} \mid (N_{m-1} = n)$, as in (S-8.4),

$$\mathbb{E}[X] = (1-p) \cdot \mathbb{E}[\exp\{t(N_m + n + 1)\}] + p \cdot \exp\{t(n+1)\}$$

or $\mathbb{E}[X] = (1-p)\, e^{tn} e^t M_m + p e^{tn} e^t$. Taking expectations w.r.t. N_{m-1} gives, with $q = 1-p$,

$$M_m = q\, e^t M_m M_{m-1} + p\, e^t M_{m-1} \quad \text{or} \quad M_m = \frac{p\, e^t M_{m-1}}{1 - q\, e^t M_{m-1}}.$$

Solution to Problem 8.14: After placing the first cut, you are left with a line of unit length. Denote the position of the second cut, i.e., its distance from the left endpoint of the line, by X, and note that $X \sim \mathrm{Unif}(0,1)$. The position of the red dot, say Y, has a uniform distribution on the unit interval as well, independently of X.

Define a random variable Z as the length of the piece that contains the dot. We can write

$$\begin{aligned} Z &= X\mathbb{I}_{(X>Y)} + (1-X)\mathbb{I}_{(X<Y)} \\ &= X\mathbb{I}_{(X>Y)} + (1-X)\mathbb{I}_{(1-X\ >\ 1-Y)}. \end{aligned}$$

Taking expectations,

$$\mathbb{E}[Z] = \mathbb{E}[X \mid X > Y]\Pr(X > Y) + \mathbb{E}[1-X \mid 1-X > 1-Y]\Pr(1-X > 1-Y).$$

As both $1-X$ and $1-Y$ have uniform distributions, the two terms on the right hand side are equal, and using $\Pr(X>Y) = 1/2$, we have

$$\mathbb{E}[Z] = \mathbb{E}[X \mid X > Y] = \int_0^1 x f_{X|X>Y}(x)\ dx.$$

Now, observe that, in general,

$$F_{X|X>Y}(t) = \frac{\Pr(X < t \wedge X > Y)}{\Pr(X > Y)} = \frac{\int_{-\infty}^{t}\int_{-\infty}^{x} f_{X,Y}(x,y)\ dy\ dx}{\int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{X,Y}(x,y)\ dy\ dx} =: \frac{I(t)}{I(\infty)}.$$

Then, similar to the derivation of (8.48),

$$I(t) = \frac{1}{2}[F_X(t)]^2,$$

or

$$F_{X|X>Y}(t) = \frac{I(t)}{I(\infty)} = [F_X(t)]^2,$$

so that

$$f_{X|X>Y}(t) = \frac{d}{dt}F_{X|X>Y}(t) = 2F_X(t)f_X(t).$$

For the case at hand,

$$f_{X|X>Y}(t) = 2t,$$

and thus

$$\mathbb{E}[Z] = \int_0^1 x f_{X|X>Y}(x)\ dx = \int_0^1 2x^2 dx = \left[\frac{2}{3}x^3\right]_0^1 = \frac{2}{3}.$$

This result is easily confirmed by simulation by using the following Matlab code:

```
x=rand(10000,1); y=rand(10000,1); z=(y<x).*x + (y>x).*(1-x); mean(z)
```

Solutions to Chapter 9: Multivariate Transformations

Solution to Problem 9.1: This is a special case of the derivation in Example 9.11 with $n = 1$. Let $r = x/(x+y)$ and $s = x + y$, so that $x = rs$ and $y = s(1-r)$. Then

$$\mathbf{J} = \begin{pmatrix} \partial x/\partial r & \partial x/\partial s \\ \partial y/\partial r & \partial y/\partial s \end{pmatrix} = \begin{pmatrix} s & r \\ -s & 1-r \end{pmatrix}, \quad \det \mathbf{J} = s$$

and

$$f_{R,S}(r,s) = |s| \, f_{X,Y}(rs, s(1-r)) \quad \text{and} \quad f_R(r) = \int_{-\infty}^{\infty} |s| \, f_{X,Y}(rs, s(1-r)) \, ds.$$

a) For $X, Y \overset{\text{iid}}{\sim} \text{Gam}(a,b)$,

$$f_{X,Y}(x,y) = \frac{b^a}{\Gamma(a)} \frac{b^a}{\Gamma(a)} x^{a-1} y^{a-1} \exp(-bx - by) \, \mathbb{I}_{(0,\infty)}(x) \, \mathbb{I}_{(0,\infty)}(y)$$

and, temporarily omitting the indicator functions,

$$\begin{aligned} f_R(r) &= \int_{-\infty}^{\infty} |s| \, f_{X,Y}(rs, s(1-r)) \, ds \\ &= \frac{b^a}{\Gamma(a)} \frac{b^a}{\Gamma(a)} \int_0^{\infty} s \, (rs)^{a-1} (s(1-r))^{a-1} \exp(-b(rs) - b(s(1-r))) \, ds \\ &= \frac{b^a}{\Gamma(a)} \frac{b^a}{\Gamma(a)} r^{a-1} (1-r)^{a-1} \int_0^{\infty} s^{2a-1} \exp(-bs) \, ds \\ &= \frac{\Gamma(2a)}{\Gamma(a)\Gamma(a)} r^{a-1} (1-r)^{a-1}, \end{aligned}$$

where

$$\int_0^{\infty} s^{2a-1} \exp(-bs) \, ds = \frac{\Gamma(2a)}{b^{2a}},$$

which follows directly from the gamma density. Regarding the indicator functions, note that $\mathbb{I}_{(0,\infty)}(rs) \, \mathbb{I}_{(0,\infty)}(s(1-r))$ implies that r and $1-r$ have the same sign, which implies that r lies between 0 and 1. Thus

$$f_R(r) = \frac{\Gamma(2a)}{\Gamma(a)\Gamma(a)} r^{a-1} (1-r)^{a-1} \, \mathbb{I}_{(0,1)}(r),$$

i.e., $R \sim \text{Beta}(a,a)$, as was found in Example 9.11.

b) For $X, Y \overset{\text{iid}}{\sim} \text{N}(0,1)$, let $k = 2r^2 + 1 - 2r$ and $u = ks^2/2$ so that

$$\begin{aligned} f_R(r) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} |s| \exp\left(-\frac{1}{2}\left((rs)^2 + (s(1-r))^2\right)\right) ds \\ &= \frac{1}{\pi} \int_0^{\infty} s \exp\left(-\frac{k}{2} s^2\right) ds \\ &= \frac{1}{\pi(2r^2 - 2r + 1)} = \frac{1}{\pi} \frac{1}{2} \frac{1}{1 + \left(\frac{r-1/2}{1/2}\right)^2}, \end{aligned}$$

i.e., $R \sim \text{Cau}(1/2, 1/2)$, a location-scale Cauchy random variable.

Solution to Problem 9.2:

a) Observe that if w_1 were 0 and w_2 were 1, we would have a special case of Example 9.11.

b) With

$$\begin{aligned} f_{X_1,X_2}(x_1,x_2) &= f_{X_1}(x_1) f_{X_2}(x_2) \\ &= \frac{1}{2\pi} x_1^{-\frac{1}{2}} x_2^{-\frac{1}{2}} e^{-\frac{1}{2}(x_1+x_2)} \mathbb{I}_{(0,\infty)}(x_1) \mathbb{I}_{(0,\infty)}(x_2), \end{aligned}$$

define

$$y_1 = \frac{w_1 x_1 + w_2 x_2}{x_1 + x_2} \quad \text{and} \quad y_2 = x_1 + x_2,$$

which yields

$$x_2 = y_2 \frac{y_1 - w_1}{w_2 - w_1} \quad \text{and} \quad x_1 = y_2 \frac{w_2 - y_1}{w_2 - w_1},$$

so that

$$\det \mathbf{J} = \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_2}{\partial y_1} \\ \dfrac{\partial x_1}{\partial y_2} & \dfrac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} -\dfrac{y_2}{w_2 - w_1} & \dfrac{y_2}{w_2 - w_1} \\ \dfrac{w_2 - y_1}{w_2 - w_1} & \dfrac{y_1 - w_1}{w_2 - w_1} \end{vmatrix} = \frac{y_2}{w_1 - w_2}.$$

As $w_2 > w_1$ and $y_2 > 0$, taking absolute values yields $\left| \dfrac{y_2}{w_1 - w_2} \right| = \dfrac{y_2}{w_2 - w_1}$.

Thus, without the indicator functions, $f_{Y_1,Y_2}(y_1, y_2)$ is

$$\begin{aligned} & |\det \mathbf{J}| \cdot f_{X_1,X_2}(x_1,x_2) \\ = \; & \frac{y_2}{w_2 - w_1} \cdot \frac{1}{2\pi} (w_2 - w_1)(y_1 - w_1)^{-1/2} (w_2 - y_1)^{-1/2} y_2^{-1/2} y_2^{-1/2} e^{-y_2/2} \\ = \; & \frac{1}{\pi} (y_1 - w_1)^{-1/2} (w_2 - y_1)^{-1/2} \cdot \frac{1}{2} e^{-y_2/2}. \end{aligned}$$

For the indicator functions, using the fact that, by assumption, $w_2 - w_1 > 0$ and, by construction, $y_2 = x_1 + x_2$ is positive,

$$\mathbb{I}_{(0,\infty)}\left(y_2 \frac{w_2 - y_1}{w_2 - w_1} \right) \mathbb{I}_{(0,\infty)}\left(y_2 \frac{y_1 - w_1}{w_2 - w_1} \right) = \mathbb{I}_{(0,\infty)}(w_2 - y_1) \mathbb{I}_{(0,\infty)}(y_1 - w_1) \mathbb{I}_{(0,\infty)}(y_2),$$

which reduces to

$$\mathbb{I}_{(0,w_2)}(y_1) \mathbb{I}_{(w_1,\infty)}(y_1) \mathbb{I}_{(0,\infty)}(y_2) = \mathbb{I}_{(w_1,w_2)}(y_1) \mathbb{I}_{(0,\infty)}(y_2)$$

because, also by assumption, $w_1 > 0$, $w_2 > 0$ and, by construction, $y_1 > 0$. That is,

$$f_{Y_1,Y_2}(y_1,y_2) = \frac{1}{\pi} (y_1 - w_1)^{-1/2} (w_2 - y_1)^{-1/2} \mathbb{I}_{(w_1,w_2)}(y_1) \; \frac{1}{2} e^{-y_2/2} \mathbb{I}_{(0,\infty)}(y_2).$$

The indicator functions could also be obtained with less "formality". From its definition, Y_2 is clearly positive. For Y_1, taking the extreme case of $X_1 = 0$ or $X_2 = \infty$ (i.e., $X_1/X_2 \to 0$), which yields $Y_1 = w_2$, while $X_2 = 0$ or $X_1 = \infty$ (i.e., $X_2/X_1 \to 0$) yields $Y_1 = w_1$. That Y_1 cannot exceed w_2 is seen by noting that $Y_1 > w_2$ would imply $w_1 X_1 + w_2 X_2 > w_2 X_1 + w_2 X_2$, which is impossible, given that $w_1 < w_2$ and $X_1 > 0$. Similar reasoning shows that $Y_1 > w_1$. Thus, we obtain $\mathbb{I}_{(w_1,w_2)}(y_1)$.

Because of the product representation, Y_1 and Y_2 are independent. Also, we recognize the latter density as a χ^2 with 2 degrees of freedom, as was expected from the results in Example 9.11 and the relation between χ^2 and gamma random variables.

c) Using the suggested substitution, $y_1 = (w_2 - w_1)\,u + w_1$, $dy_1 = (w_2 - w_1)\,du$, $y_1 = w_1 \Longleftrightarrow u = 0$ and $y_1 = w_2 \Longleftrightarrow u = 1$ and, thus,

$$\begin{aligned}
& \frac{1}{\pi}\int_{w_1}^{w_2} (y_1 - w_1)^{-1/2}\,(w_2 - y_1)^{-1/2}\,dy_1 \\
= & \frac{1}{\pi}\int_{0}^{1} (1-u)^{-1/2}\,(w_2 - w_1)^{-1/2}\,u^{-1/2}\,(w_2 - w_1)^{-1/2}\cdot(w_2 - w_1)\,du \\
= & \frac{1}{\pi}\,\text{Beta}\left(\frac{1}{2},\frac{1}{2}\right) = 1.
\end{aligned}$$

Solution to Problem 9.3: The inverse transformation is

$$\begin{aligned}
X_1 &= Y_1, \\
X_2 &= Y_2 - Y_1, \\
X_3 &= Y_3 - Y_2, \\
&\vdots \\
X_n &= Y_n - Y_{n-1},
\end{aligned}$$

with inverse Jacobian

$$\mathbf{J}^{-1} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_n}{\partial x_2} \\ \vdots & \vdots & \ddots & \\ \frac{\partial y_1}{\partial x_n} & \frac{\partial y_2}{\partial x_n} & \cdots & \frac{\partial y_n}{\partial x_n} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

and $\det(\mathbf{J}) = 1$, so that $f_{\mathbf{Y}}(\mathbf{y})$ is

$$\begin{aligned}
& f_{\mathbf{X}}(\mathbf{x}) \\
= & \lambda^n e^{-\lambda(y_1)} e^{-\lambda(y_2 - y_1)} \cdots e^{-\lambda(s - y_{n-1})} \mathbb{I}_{(0,\infty)}(y_1)\,\mathbb{I}_{(0,\infty)}(y_2 - y_1)\cdots\mathbb{I}_{(0,\infty)}(s - y_{n-1}) \\
= & \lambda^n e^{-\lambda s} \mathbb{I}_{(0,\infty)}(y_1)\,\mathbb{I}_{(0,\infty)}(y_2 - y_1)\cdots\mathbb{I}_{(0,\infty)}(s - y_{n-1}).
\end{aligned}$$

Integrating out $Y_{n-1}, Y_{n-2}, \ldots, Y_1$, we obtain

$$\begin{aligned}
f_S(s) &= \lambda^n e^{-\lambda s}\int_0^s\int_0^{y_{n-1}}\cdots\int_0^{y_5}\int_0^{y_4}\int_0^{y_3}\int_0^{y_2} dy_1\,dy_2 dy_3 dy_4\cdots dy_{n-2} dy_{n-1}\mathbb{I}_{(0,\infty)}(s) \\
&= \lambda^n e^{-\lambda s}\int_0^s\int_0^{y_{n-1}}\cdots\int_0^{y_5}\int_0^{y_4}\int_0^{y_3} y_2 dy_2\,dy_3 dy_4\cdots dy_{n-2} dy_{n-1}\mathbb{I}_{(0,\infty)}(s) \\
&= \frac{1}{2}\lambda^n e^{-\lambda s}\int_0^s\int_0^{y_{n-1}}\cdots\int_0^{y_5}\int_0^{y_4} y_3^2 dy_3\,dy_4\,\cdots dy_{n-2} dy_{n-1}\mathbb{I}_{(0,\infty)}(s) \\
&= \frac{1}{2}\frac{1}{3}\lambda^n e^{-\lambda s}\int_0^s\int_0^{y_{n-1}}\cdots\int_0^{y_5} y_4^3\,dy_4\,\cdots dy_{n-2} dy_{n-1}\mathbb{I}_{(0,\infty)}(s) \\
&\;\;\vdots \\
&= \frac{\lambda^n}{(n-1)!} e^{-\lambda s} s^{n-1}\mathbb{I}_{(0,\infty)}(s).
\end{aligned}$$

Solution to Problem 9.4:

a) $f_{Y_1,Y_2,Y_3}(y_1,y_2,y_3) = |\mathbf{J}|\, f_{X_1,X_2,X_3}(x_1,x_2,x_3)$, where

$$\mathbf{J}^{-1} = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_2}{\partial x_1} & \frac{\partial g_3}{\partial x_1} \\ \frac{\partial g_1}{\partial x_2} & \frac{\partial g_2}{\partial x_2} & \frac{\partial g_3}{\partial x_2} \\ \frac{\partial g_1}{\partial x_3} & \frac{\partial g_2}{\partial x_3} & \frac{\partial g_3}{\partial x_3} \end{bmatrix} := \left[\frac{\partial g_i}{\partial x_j}\right]_{i,j=1,2,3} = \begin{bmatrix} \frac{x_2+x_3}{s^2} & \frac{x_3}{s^2} & 1 \\ -\frac{x_1}{s^2} & \frac{x_3}{s^2} & 1 \\ -\frac{x_1}{s^2} & -\frac{x_1+x_2}{s^2} & 1 \end{bmatrix}$$

and

$$g_1 = \frac{x_1}{s}, \quad g_2 = \frac{x_1+x_2}{s} \quad \text{and } g_3 = s,$$

where $s = x_1 + x_2 + x_3$. A little work shows

$$\left|\det \mathbf{J}^{-1}\right| = \frac{(x_1+x_2+x_3)^2}{s^4} = s^{-2}.$$

Rewriting in terms of the y_i,

$$x_1 = y_1 y_3, \quad x_2 = y_3(y_2 - y_1) \quad \text{and} \quad x_3 = y_3(1-y_2)$$

so that

$$\begin{aligned} |\det \mathbf{J}|\, f_{X_1,X_2,X_3}(x_1,x_2,x_3) &= s^2 e^{-(y_3(1-y_2)+y_3(y_2-y_1)+y_1y_3)} \\ &= y_3^2 e^{-y_3}\mathbb{I}_{(0,\infty)}(y_3), \qquad 0 < y_1 < y_2 < 1, \end{aligned}$$

where the ranges follow because $X_i > 0$, $i = 1,2,3$.

b) From Example 9.11,

$$X_2 + X_3 \sim \text{Gam}(r=2, \lambda=1), \text{independent of } X_1 \sim \text{Gam}(r=1,\lambda=1)$$

and

$$Y_1 = \frac{X_1}{X_1 + (X_2+X_3)} \sim \text{Beta}(1,2).$$

Similarly, $Y_2 \sim \text{Beta}(2,1)$ and $Y_3 \sim \text{Gam}(r=3,\lambda=1)$. Alternatively,

$$f_{Y_1}(y_1) = \int_0^\infty \int_{y_1}^1 y_3^2\, e^{-y_3} dy_2 dy_3 = 2(1-y_1)\mathbb{I}_{(0,1)}(y_1)$$

which is the Beta $(1,2)$ pdf. The integral $\int_0^\infty y^2\, e^{-y} dy$ can be evaluated either by integration by parts, i.e., $\int u dv = uv - \int v du$ with $u = y^2$ and $dv = e^{-y}dy$, or as follows. Notice that the integral is precisely $\mathbb{E}\left[Y^2\right] = \mu_2'$, the second raw moment, with $Y \sim \text{Exp}(1)$. Because, in general, $\mu_2' = \mu^2 + \sigma^2$, it follows that $\mu_2' = 1^2 + 1 = 2$, recalling that the mean and variance of an exponential distribution.

Similarly,

$$f_{Y_2}(y_2) = \int_0^\infty \int_0^{y_2} y_3^2\, e^{-y_3} dy_1 dy_3 = 2y_2\mathbb{I}_{(0,1)}(y_2)$$

which is the Beta $(2,1)$ pdf. Finally,

$$f_{Y_3}(y_3) = \int_0^1 \int_0^{y_2} y_3^2\, e^{-y_3} dy_1 dy_2 = \frac{1}{2} y_3^2\, e^{-y_3}\mathbb{I}_{(0,\infty)}(y_3)$$

which we recognize as the Gam $(r=3,\lambda=1)$ pdf.

c)

$$f_{Y_1,Y_2}(y_1,y_2) = \int_0^\infty y_3^2 e^{-y_3} dy_3 = 2 \qquad 0 \le y_1 \le y_2 \le 1,$$

which is a triangle with corners $(0,0)$, $(0,1)$ and $(1,1)$, with area $\frac{1}{2}$. Clearly,

$$f_{Y_1,Y_2,Y_3}(y_1,y_2,y_3) = f_{Y_1,Y_2}(y_1,y_2) \cdot f_{Y_3}(y_3)$$

so that they are independent.

d) $f_{Y_1|Y_2}(y_1 \mid y_2) = \dfrac{f_{Y_1,Y_2}(y_1,y_2)}{f_{Y_2}(y_2)} = y_2^{-1} \mathbb{I}_{(0,1)}(y_2) \mathbb{I}_{(0,y_2)}(y_1)$, which is uniform over the interval $0 < y_1 < y_2$.

Solution to Problem 9.5:

a) With $W = X$, the bivariate transformation $\{X,Y\} \to \{W,Z\}$ yields $X = W$, $Y = Z - W$ and $|J| = 1$, so that

$$f_{W,Z}(w,z) = f_{X,Y}(x,y) = \lambda^2 e^{-\lambda(x+y)} = \lambda^2 e^{-\lambda z} \mathbb{I}_{(0,z)}(w) \mathbb{I}_{(0,\infty)}(z).$$

Thus, $f_Z(z) = f_{X+Y}(z) = \int_0^z f_{W,Z}(w,z)\, dw = z\lambda^2 e^{-\lambda z} \mathbb{I}_{(0,\infty)}(z)$, or the gamma density with shape 2 and scale λ^{-1}.

b)

$$f_{X|Z=z}(x \mid z) = f_{W|Z=z}(w \mid z) = \frac{f_{W,Z}(w,z)}{f_Z(z)} = z^{-1} \mathbb{I}_{(0,z)}(x).$$

c) From the previous question,

$$\Pr\left(\frac{Z}{3} < X < \frac{Z}{2} \mid Z = z\right) = \int_{z/3}^{z/2} f_{X|Z=z}(x \mid z)\, dx = \frac{1}{6}.$$

d) Using the given formula,

$$\begin{aligned} f_D(d) &= \int_{-\infty}^{\infty} f_X(d+y) f_Y(y)\, dy \\ &= \lambda^2 \int_{-\infty}^{\infty} e^{-\lambda(d+y)} e^{-\lambda y} \mathbb{I}_{(0,\infty)}(d+y) \mathbb{I}_{(0,\infty)}(y)\, dy, \end{aligned}$$

but

$$\mathbb{I}_{(0,\infty)}(d+y) \mathbb{I}_{(0,\infty)}(y) \Rightarrow -d < y,\ y > 0 \Rightarrow \begin{cases} 0 < y, & \text{for } d > 0, \\ -d < y, & \text{for } d < 0, \end{cases}$$

so that

$$\begin{aligned} f_D(d) &= \mathbb{I}_{(-\infty,0)}(d) \int_{-d}^{\infty} \lambda^2 e^{-\lambda(d+y)} e^{-\lambda y} dy + \mathbb{I}_{(0,\infty)}(d) \int_0^\infty \lambda^2 e^{-\lambda(d+y)} e^{-\lambda y} dy \\ &= \frac{1}{2}\lambda e^{\lambda d} \mathbb{I}_{(-\infty,0)}(d) + \frac{1}{2}\lambda e^{-\lambda d} \mathbb{I}_{(0,\infty)}(d) = \frac{1}{2}\lambda e^{-\lambda|d|}, \end{aligned}$$

similar to Example 9.5.

e) As $D \sim \text{Lap}(\lambda)$, the density of $W = |D|$ should be just twice the density of D and with support the positive real line instead of the whole real line. That is, $W \sim \text{Exp}(\lambda)$. More formally, the cdf of W is given by

$$F_W(w) = \Pr(W \le w) = \Pr(-w \le D \le w) = F_D(w) - F_D(-w),$$

and differentiating,

$$\begin{aligned} f_W(w) &= f_D(w) + f_D(-w) = \frac{1}{2}\lambda e^{-\lambda|w|} + \frac{1}{2}\lambda e^{-\lambda|-w|} \\ &= \lambda e^{-\lambda w} \mathbb{I}_{(0,\infty)}(w) \end{aligned}$$

with expected value λ^{-1}.

f) Let $X_{(1)} = \min(X,Y)$, $X_{(2)} = \max(X,Y)$ and $R = X_{(2)} - X_{(1)}$ denote the range of X and Y. Then $f_{R|X_{(1)}=a}(r \mid a) = f_{R,X_{(1)}}(r,a) \,/\, f_{X_{(1)}}(a)$ with the joint density $f_{R,X_{(1)}}(r,a)$ obtained from the bivariate transformation $g_1 = x_{(1)}$, $g_2 = x_{(2)} - x_{(1)}$, $\det \mathbf{J} = \left|\begin{smallmatrix} 1 & 0 \\ -1 & 1 \end{smallmatrix}\right| = 1$ so that $f_{R,X_{(1)}}(r,a) = f_{X_{(1)},X_{(2)}}(a, a+r) = 2!\lambda e^{-\lambda a}\lambda e^{-\lambda(a+r)}$ and, differentiating (**??**), $f_{X_{(1)}}(a) = n\left[1 - F_X(a)\right]^{n-1} f_X(a)$ with $n = 2$ or

$$f_{R|X_{(1)}=a}(r \mid a) = \frac{f_{R,X_{(1)}}(r,a)}{f_{X_{(1)}}(a)} = \frac{2\lambda e^{-\lambda a}\lambda e^{-\lambda(a+r)}}{2\lambda e^{-\lambda a} e^{-\lambda a}} = \lambda e^{-\lambda r}\mathbb{I}_{(0,\infty)}(r).$$

Solution to Problem 9.6:

a) Let $n = 2$, so that $X_2 = g_2^{-1}(\mathbf{Y}) = Y_2$ and $X_1 = g_1^{-1}(\mathbf{Y}) = \pm\sqrt{Y_1 - Y_2^2}$. Splitting this into two regions,

$$f_{\mathbf{Y}}(\mathbf{y}) = \left|\det \mathbf{J}_1\right| \; f_{\mathbf{X}}(x_1,x_2)\,\mathbb{I}_{(-\infty,0)}(x_1) + \left|\det \mathbf{J}_2\right| \; f_{\mathbf{X}}(x_1,x_2)\,\mathbb{I}_{(0,-\infty)}(x_1),$$

where

$$\mathbf{J}_1 = \begin{bmatrix} \partial x_1/\partial y_1 & \partial x_1/\partial y_2 \\ \partial x_2/\partial y_1 & \partial x_2/\partial y_2 \end{bmatrix} = \begin{bmatrix} -\left(y_1 - y_2^2\right)^{-1/2}/2 & y_2\left(y_1 - y_2^2\right)^{-1/2} \\ 0 & 1 \end{bmatrix}$$

and similarly for $\mathbf{J}_2$, and, in both cases, $|\mathbf{J}_i| = \left(y_1 - y_2^2\right)^{-1/2}/2$. Thus,

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{2}\left(y_1 - y_2^2\right)^{-1/2}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(y_1 - y_2^2\right)}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y_2^2} + \text{same}$$

or

$$f_{\mathbf{Y}}(\mathbf{y}) = \left(y_1 - y_2^2\right)^{-1/2}\frac{1}{2\pi}e^{-\frac{1}{2}y_1}\mathbb{I}_{\left(y_2^2,\infty\right)}(y_1),$$

where the indicator function follows from

$$Y_1 = X_1^2 + X_2^2 \quad \text{and} \quad 0 \le X_2^2 \le X_1^2 + X_2^2$$

or $0 \le Y_2^2 \le Y_1$. This also implies $-\sqrt{Y_1} \le Y_2 \le \sqrt{Y_1}$, from which we have

$$f_{Y_1}(y_1) = \int_{-\sqrt{y_1}}^{\sqrt{y_1}} f_{\mathbf{Y}}(\mathbf{y})\,dy_2 = \frac{1}{2\pi}e^{-\frac{1}{2}y_1}\int_{-\sqrt{y_1}}^{\sqrt{y_1}}\left(y_1 - y_2^2\right)^{-1/2}dy_2.$$

From the hint,

$$\int \frac{1}{\sqrt{y_1 - y_2^2}}dy_2 = c + \arcsin\left(\frac{y_2}{\sqrt{y_1}}\right),$$

and, clearly, $\arcsin(1) = \pi/2$ and $\arcsin(-1) = -\pi/2$, so that

$$f_{Y_1}(y_1) = \frac{1}{2}\exp(-y_1/2)\,\mathbb{I}_{(0,\infty)}(y_1),$$

where the indicator follows from the definition of $Y_1 = \sum_{i=1}^2 X_i^2$. Thus, as in Example 9.12, $Y_1 \sim \chi_2^2$.

b) This integral is, using $v = (y_0 - u)/y_0$, $u = (1 - v)\,y_0$, $du = -y_0 dv$,

$$\begin{aligned} J &= \int_0^{y_0} u^{m/2-1}(y_0 - u)^{-1/2}du = -\int_1^0 \left((1-v)\,y_0\right)^{m/2-1}(vy_0)^{-1/2}y_0 dv \\ &= y_0^{(m-1)/2}\int_0^1 v^{(1/2)-1}(1-v)^{m/2-1}dv \\ &= y_0^{(m-1)/2}B\left(\frac{1}{2},\frac{m}{2}\right) = y_0^{(m-1)/2}\frac{\Gamma(1/2)\,\Gamma(m/2)}{\Gamma((m+1)/2)}. \end{aligned}$$

c) From the result using $n = 2$, one might guess that the general case might lead to $Y_1 \sim \chi^2_n$, which is true. Now $X_i = g_i^{-1}(\mathbf{Y}) = Y_i$, $i = 2, \ldots, n$ and $X_1 = g_1^{-1}(\mathbf{Y}) = \pm\sqrt{Y_1 - \sum_{i=2}^n Y_i^2}$. We again need to split the support of $\mathbf{X}$ into two regions as before, but we have seen for the $n = 2$ case that the components are the same. Thus,

$$f_{\mathbf{Y}}(\mathbf{y}) = 2\left|\mathbf{J}^{-1}\right|^{-1} f_{\mathbf{X}}(\mathbf{x}),$$

where we use $\mathbf{J}^{-1}$ instead of $\mathbf{J}$ because it is algebraically more convenient. We have

$$\mathbf{J}^{-1} = \begin{bmatrix} \partial y_1/\partial x_1 & \partial y_1/\partial x_2 & \cdots & \partial y_1/\partial x_n \\ \partial y_2/\partial x_1 & \partial y_2/\partial x_2 & \cdots & \partial y_2/\partial x_n \\ \vdots & \vdots & \ddots & \vdots \\ \partial y_n/\partial x_1 & \partial y_n/\partial x_2 & \cdots & \partial y_n/\partial x_n \end{bmatrix} = \begin{bmatrix} 2x_1 & 2x_2 & \cdots & 2x_n \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix},$$

with determinant $\left|\mathbf{J}^{-1}\right| = 2x_1 = 2\sqrt{Y_1 - \sum_{i=2}^n Y_i^2}$. Then, defining

$$D = y_1 - \sum_{i=2}^n y_i^2$$

for convenience, $\left|\mathbf{J}^{-1}\right|^{-1} = D^{-1/2}/2$ and

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= 2 \cdot \frac{1}{2} D^{-1/2} \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\left(D + \sum_{i=2}^n y_i^2\right)\right\} \mathbb{I}_{\mathcal{S}_{\mathbf{Y}}}(\mathbf{y}) \\ &= D^{-1/2} \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}y_1} \mathbb{I}_{\mathcal{S}_{\mathbf{Y}}}(\mathbf{y}). \end{aligned} \tag{S-9.1}$$

(To check, with $n = 2$ and no indicator functions, this reduces to

$$f_{Y_1,Y_2}(y_1, y_2) = \left(y_1 - y_2^2\right)^{-1/2} \frac{1}{2\pi} e^{-\frac{1}{2}y_1}$$

which agrees with the direct derivation above.)
Inserting D into (S-9.1) and setting up the integral,

$$f_{Y_1}(y_1) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}y_1} \int \cdots \int_{\mathcal{S}} \left(y_1 - \sum_{i=2}^n y_i^2\right)^{-\frac{1}{2}} dy_2 \cdots dy_n, \tag{S-9.2}$$

where

$$\mathcal{S} = \left\{(y_2, \ldots, y_n) \in \mathbb{R}^{n-1} : 0 < \sum_{i=2}^n y_i^2 < y_1\right\}.$$

We wish to apply Liouville's result to the integral

$$I = \int \cdots \int_{\mathcal{S}} \left(y_1 - \sum_{i=2}^n y_i^2\right)^{-\frac{1}{2}} dy_2 \cdots dy_n,$$

which we rewrite as

$$I = \int \cdots \int_{\mathcal{S}} \left(y_0 - \sum_{i=1}^m y_i^2\right)^{-\frac{1}{2}} dy_1 \cdots dy_m,$$

where

$$\mathcal{S} = \left\{(y_1, \ldots, y_m) \in \mathbb{R}^m : 0 < \sum_{i=1}^m y_i^2 < y_0\right\}$$

and $m = n - 1$. This is almost in the form of (9.13) when taking $p_i = 2$ and $a_i = b_i = 1$, $i = 1, \ldots, m$, (so that $r_i = 1/2$ and $R = m/2$) as well as $t_1 = 0$, $t_2 = y_0$ and $f(u) = (y_0 - u)^{-1/2}$.

The problem is that the condition $x_i \geq 0$ in (9.13) is not fulfilled. However, what we can compute is

$$I' = \int \cdots \int_{\mathcal{S}'} \left(y_0 - \sum_{i=1}^{m} y_i^2 \right)^{-\frac{1}{2}} dy_1 \cdots dy_m,$$

where

$$\mathcal{S}' = \left\{ (y_1, \ldots, y_m) \in \mathbb{R}_+^m : 0 < \sum_{i=1}^{m} y_i^2 < y_0 \right\},$$

i.e., each y_i is restricted to be positive. Then, via the symmetry of the standard normal distribution about zero and the fact that each y_i enters the function f as y_i^2, we see that $I = 2^m I'$. Now, using (9.13),

$$I = 2^m I' = 2^m \frac{(1/2)^m \pi^{m/2}}{\Gamma(m/2)} \int_0^{y_0} u^{m/2-1} (y_0 - u)^{-1/2} \, du = \frac{\pi^{m/2}}{\Gamma(m/2)} J,$$

where the integral J was shown above to be

$$J = y_0^{(m-1)/2} \frac{\Gamma(1/2)\Gamma(m/2)}{\Gamma((m+1)/2)}.$$

Then, recalling that $m = n - 1$,

$$I = \frac{\pi^{m/2}}{\Gamma(m/2)} y_0^{(m-1)/2} \frac{\Gamma(1/2)\Gamma(m/2)}{\Gamma((m+1)/2)} = \frac{\pi^{n/2}}{\Gamma(n/2)} y_0^{n/2-1}.$$

Renaming y_0 back to y_1, (S-9.2) gives

$$f_{Y_1}(y_1) = \frac{e^{-y_1/2}}{(2\pi)^{n/2}} \frac{\pi^{n/2}}{\Gamma(n/2)} y_1^{n/2-1} = \frac{1}{2^{n/2}\Gamma(n/2)} e^{-y_1/2} y_1^{n/2-1},$$

which is the χ_n^2 density.

Solution to Problem 9.7:

a) It follows directly from (A.166) and (9.1) that $f_{R,\Theta}(r, \theta) = r f_{X,Y}(x, y)$, which yields (9.14).

b) From (7.65), this is

$$f_S(s) = f_R(r) \left| \frac{dr}{ds} \right| = r e^{-r^2/2} \left(\frac{1}{2} s^{-1/2} \right) \mathbb{I}_{(0,\infty)} \left(s^{1/2} \right) = \frac{1}{2} e^{-s/2} \mathbb{I}_{(0,\infty)}(s),$$

so that $S \sim \text{Exp}(1/2)$.

c) For positive u,

$$\begin{aligned} \Pr(-2 \ln U_1 < u) &= \Pr(\ln U_1 > -u/2) = \Pr(U_1 > \exp(-u/2)) \\ &= 1 - \exp(-u/2), \end{aligned}$$

which is the survivor function of an $\text{Exp}(1/2)$ random variable.

d) As the inverse of the polar coordinate transformation is $X = R\cos\Theta$ and $Y = R\sin\Theta$, two iid $\text{N}(0, 1)$ r.v.s can be generated by setting

$$X = R\cos\Theta = \sqrt{R^2}\cos\Theta = \sqrt{-2\ln U_1}\cos(2\pi U_2)$$

and similarly for Y, which yield (9.9).

Solution to Problem 9.8: From Example 9.3, $S \sim \text{Gam}(n, \lambda)$. We now show that $S \mid (X_1 = x)$ is a location shifted gamma random variable: With $Z = \sum_{i=2}^{n} X_i \sim \text{Gam}(n-1, \lambda)$ and $Y = (S \mid X_1) = (Z + x)$, a simple transformation gives

$$f_Y(y) = f_Z(z) \frac{dz}{dy} = \frac{\lambda^{n-1}}{\Gamma(n-1)} (y-x)^{(n-1)-1} e^{-\lambda(y-x)}.$$

From Bayes' theorem and the fact that $X_1 \leq S$, we obtain

$$\begin{aligned} f_{X_1|S}(x \mid s) &= \frac{f_{S|X_1}(s \mid x) f_{X_1}(x)}{f_S(s)} = \frac{\frac{\lambda^{n-1}}{\Gamma(n-1)} (s-x)^{n-2} e^{-\lambda(s-x)} \lambda e^{-\lambda x}}{\frac{\lambda^n}{\Gamma(n)} s^{n-1} e^{-\lambda s}} \\ &= \frac{n-1}{s^{n-1}} (s-x)^{n-2} \mathbb{I}_{(0,s)}(x), \quad n > 1. \end{aligned}$$

For the expected value, use substitution $u = s - x$ to get

$$\begin{aligned} \mathbb{E}[X_1 \mid S = s] &= \frac{n-1}{s^{n-1}} \int_0^s x (s-x)^{n-2} dx = -\frac{n-1}{s^{n-1}} \int_s^0 (s-u) u^{n-2} du \\ &= \frac{n-1}{s^{n-1}} s \int_0^s u^{n-2} du - \frac{n-1}{s^{n-1}} \int_0^s u^{n-1} du = s - \frac{n-1}{s^{n-1}} \frac{s^n}{n} = \frac{s}{n}. \end{aligned}$$

Similarly, the distribution function is

$$\begin{aligned} F_{X_1|S}(t) &= \int_0^t f_{X_1|S}(x \mid s) \mathbb{I}_{(0,s)}(t) dx + \mathbb{I}_{[s,\infty)}(t) \\ &= 1 - \frac{(s-t)^{n-1}}{s^{n-1}} \mathbb{I}_{(0,s)}(t) + \mathbb{I}_{[s,\infty)}(t). \end{aligned}$$