original

SOLUTIONS AND COMMENTS FOR SELECTED EXERCISES

Chapter 1

1.1 (a) As n increases, the ratio r/n is seen to stabilize. The 'limiting frequency' definition of probability is in terms of the limits of such ratios as $n \to \infty$.

(b) See, e.g., Feller (1957, p. 84). Note the occurrence of 'long leads', i.e. once $(2r - n)$ becomes positive (or negative) it frequently stays so for many consecutive trials.

1.2 Buffon's needle is discussed further in Exercises 7.1–7.4. Estimate Prob(crossing) by:

no. of crossings/no. of trials

and solve to estimate π ($\hat{\pi}$, say).

(i) $\dfrac{2}{\hat{\pi}} = \dfrac{254}{390}$; $\hat{\pi} = 3.071$

(ii) $\dfrac{2}{\hat{\pi}} = \dfrac{638}{960}$; $\hat{\pi} = 3.009$

1.3 Effects of increasing the traffic at a railway station, due to re-routing of trains.
Effects of instituting fast-service tills at a bank/supermarket.
Effects of promoting more lecturers to senior lecturers in universities (see Morgan and Hirsch, 1976).
The improved stability, in high winds, of high-sided vehicles with roofs with rounded edges.
Election voting patterns.

1.5 The story is taken up by McArthur *et al.* (1976) (see Fig. 8.3).

1.6 Appointment systems are now widely used in British surgeries. For a

If $\mathbf{\Sigma} = \mathbf{I}$, as in the question, $\mathbf{Z}$ has an $N(\mathbf{0}, \mathbf{AA}')$ distribution. The translation: $\tilde{\mathbf{Z}} = \mathbf{Z} + \boldsymbol{\mu}$ is readily shown to result in $N(\boldsymbol{\mu}, \mathbf{AA}')$, as required.

2.17 $$\Pr(k \text{ events}) = \int_0^\infty \lambda_1 e^{-\lambda_1 t} e^{-\lambda_2 t} \frac{(\lambda_2 t)^k}{k!}\,\mathrm{d}t \qquad \text{for } k \geq 0$$
$$= \frac{\lambda_2^k \lambda_1}{k!} \int_0^\infty e^{-(\lambda_1+\lambda_2)t} t^k\,\mathrm{d}t = \frac{\lambda_2^k \lambda_1}{k!(\lambda_1+\lambda_2)^{k+1}} \int_0^\infty e^{-\theta}\theta^k \mathrm{d}\theta$$
$$= \frac{\lambda_1 \lambda_2^k}{(\lambda_1+\lambda_2)^{k+1}} \qquad \text{for } k \geq 0$$

i.e. a geometric distribution.

2.18 $$\Pr(Y \geq n) = \sum_{k=n}^{n+m} \binom{n+m}{k} (1-\theta)^k \theta^{n+m-k}$$
$$= \sum_{k=0}^{m} \binom{n+m}{n+k} (1-\theta)^{n+k} \theta^{m-k}$$

while $\Pr(X \leq n) = \sum_{k=0}^{m} \binom{n+k-1}{k} \theta^k (1-\theta)^n$

Therefore we require

$$\sum_{k=0}^{m} \binom{n+k-1}{k} \theta^k = \sum_{k=0}^{m} \binom{n+m}{n+k} (1-\theta)^k \theta^{m-k}$$

i.e. we require

$$\binom{n+i-1}{i} = \sum_{k=m-i}^{m} \binom{n+m}{n+k} \binom{k}{m-i} (-1)^{k-m+i} \qquad \text{for } 0 \leq i \leq m,$$

and this follows simply from considering the coefficient of z^i on both sides of the identity

$$(1+z)^{n+i-1} = (1+z)^{n+m}/(1+z)^{m+1-i}$$

2.19 $e^{-n} \sum_{r=0}^{n} \frac{n^r}{r!} = \Pr(X_n \leq n)$, where X_n has a Poisson distribution of parameter n. Now (see Exercise 2.8(b)), if X_n is of this form, we can write

$$X_n = \sum_{i=1}^{n} Z_i$$

where the Z_i are independent, identically distributed Poisson random variables of parameter 1. Hence by a simple central limit theorem,

$$\Pr(X_n \leq n) \to \Phi(0) = \tfrac{1}{2} \qquad \text{as } \mathscr{E}[X_n] = n$$

Therefore $\lim_{n \to \infty} \left(e^{-n} \sum_{r=0}^{n} \frac{n^r}{r!} \right) = \frac{1}{2}$

2.22 See *ABC*, p. 389.

2.25 $y = g(x)$ does not, in this example, have a continuous derivative. Thus Equation (2.3) is not appropriate. However, we always have:

$$\Pr(Y \leq y) = \Pr(X \leq x)$$

and for $y \leq 1$,

$$\Pr(Y \leq y) = 1 - e^{-y}$$

while for $y \geq 1$,

$$\Pr(Y \leq y) = 1 - e^{-(1+y)/2}$$

2.26 See Exercise 7.25.

2.27 For $\lambda < \mu$ the queue settles down to an 'equilibrium' state. For $\lambda \geq \mu$ no such state exists, and the queue size increases ultimately without limit. See Exercise 7.24.

Chapter 3

3.1 Pr(respond 'Yes') = Pr(respond 'Yes' to (i))Pr((i) is the question)
+ Pr(respond 'Yes' to (ii))Pr((ii) is the question).

If one responds 'Yes' to (i) then one responds 'No' to (ii)
Pr(respond 'Yes' to (ii)) = 1 − Pr(respond 'Yes' to (i)) = $1 - \eta$, say
Pr(respond 'Yes') = η Pr((i)) + $(1 - \eta)$ Pr(ii)
Pr(i) and Pr(ii) are determined by a randomization device, and Pr(respond 'Yes') is estimated from the responses. For example, from a class survey of first-year mathematics students, in which group X approved of sit-ins as a form of protest (a question topical at the time), Pr(i) = 3/10; Pr(ii) = 7/10, and 34 responded 'Yes' out of 56, resulting in $\hat{\eta} = 0.38$. The students also wrote their opinion anonymously on slips which were collected, and which gave rise to $\hat{\eta} = 0.33$. See Warner (1965) for further discussion on confidence intervals, etc.

3.2 Conduct two surveys with two different probabilities of answering the innocent question, resulting in two equations in two unknowns. For further discussion, see Moors (1971). Innocent questions with known frequency of response might include month of birth, or whether an identity-number of some kind is even or odd (Campbell and Joiner, 1973).

3.3 Let θ = Pr(respondent has had an abortion)
Let proportions of balls be: p_r, p_w, p_b.
Then Pr(respond 'Yes') = $\theta p_r + p_w$.
See Greenberg *et al.* (1971).

3.4 With the question as stated in Example 3.1, 'Yes' is potentially incriminating, whereas 'No' is not. See Abdul-Ela *et al.* (1967) and Warner (1971).

3.5 In experiments of this kind, individuals typically overestimate μ, introducing a judgement bias. Here we find:

judgement: $\bar{x} = 45.56; s = 13.69$
random : $\bar{x} = 37.21; s = 10.28$.

3.6 Let X denote the recorded value.

$$\Pr(X=0) = \sum_{i=4}^{6} \Pr(\text{1st die} = i)\ \Pr(\text{2nd die} = 10-i)$$

$$= 3\cdot\tfrac{1}{6}\cdot\tfrac{1}{6} = \tfrac{1}{12} \qquad \text{(assuming independence)}$$

$$\Pr(X=1) = \sum_{i=5}^{6} \Pr(\text{1st die} = i)\ \Pr(\text{2nd die} = 11-i) = \tfrac{1}{18}.$$

For $2 \le i \le 9$,

$$\Pr(X=i) = \sum_{j=\max(1,i-6)}^{\min(6,i-1)} \Pr(\text{1st die} = j)\ \Pr(\text{2nd die} = i-j).$$

Thus $\Pr(X=2) = \frac{1}{36}$, $\Pr(X=3) = \frac{1}{18}$, $\Pr(X=4) = \frac{1}{12}$, etc.
For equiprobable random digits, the method suggested is clearly unsuitable.

3.7 $\Pr(\text{2nd coin is } H \mid \text{two tosses differ}) = \dfrac{pq}{pq+qp} = \tfrac{1}{2}$

$(p = \Pr(H) = 1-q)$.

3.8 There are 6 possibilities for both A and B, and so there are 36 possibilities in all. If we reject (i, i) results, for $1 \le i \le 6$, we get 30 possibilities, which may be used to generate uniform random digits from 0–9 as follows:

Possible outcomes from dice			Digit selected
(1, 2)	(1, 3)	(1, 4)	0
(1, 5)	(1, 6)	(2, 1)	1
(2, 3)	(2, 4)	(2, 5)	2
(2, 6)	(3, 1)	(3, 2)	3
(3, 4)	(3, 5)	(3, 6)	4
(4, 1)	(4, 2)	(4, 3)	5
(4, 5)	(4, 6)	(5, 1)	6
(5, 2)	(5, 3)	(5, 4)	7
(5, 6)	(6, 1)	(6, 2)	8
(6, 3)	(6, 4)	(6, 5)	9

Thus, for example, we choose digit 5 if we get one of (4, 1), (4, 2), (4, 3), and the conditional probability of this is:

$$(3 \times (\tfrac{1}{6}) \times (\tfrac{1}{6}))/(30/36) = \tfrac{1}{10} \qquad \text{as required.}$$

The pairs given thus result in the sequence:

$$0, 3, 1, 5, -, 9, 6, 6, 4, 0.$$

If we take these digits in fours we obtain the required numbers; e.g. 0315, 9664.

3.9 Regard HCY 7F as HCY 007F, etc. One set of 600 digits obtained in this manner is given below:

157	741	602	823	438	455	816	493	681	241
260	765	308	684	564	918	422	772	471	072
217	192	159	274	946	068	017	230	889	812
235	801	517	582	277	573	808	623	641	770
601	319	100	153	976	015	506	460	342	357
485	803	335	844	370	556	724	900	935	195
800	360	263	427	280	419	515	991	296	712
297	122	007	388	186	876	581	793	352	053
285	307	996	988	973	794	981	677	212	464
246	893	373	113	723	725	778	645	028	611

395	288	291	370	744	142	486	374	548	580
591	454	733	986	484	423	594	938	670	323
792	355	642	059	803	356	278	500	840	383
416	453	461	412	851	560	978	483	772	615
885	520	441	909	435	802	055	933	659	554
801	726	501	651	828	941	570	164	104	380
253	882	072	848	909	249	147	309	522	503
015	813	421	805	702	342	920	170	226	312
832	562	730	301	704	965	728	387	761	360
028	331	334	202	479	916	953	930	462	369

3.11 It is easy to demonstrate degeneration of this method:

(i) 55 02 00

(ii) 66 35 22 48 30 90 10 10 10

3.12 If this sequence is operated to small decimal-place accuracy then it can degenerate, as the following example shows:

0.1 0.925 0.309 0.180 0.326 0.349 0.198 0.401
0.964 0.485 0.327 0.073 0.265 0.776 0.776

3.13 We require sample mean and variance for the sample: $\{i/m,\ 0 \le i \le m-1\}$

$$\bar{x} = \frac{1}{m^2}\,\frac{m(m-1)}{2} = \frac{1}{2}\left(1 - \frac{1}{m}\right)$$

$$(m-1)s^2 = \left(\frac{1}{m^2}\sum_{i=1}^{m-1} i^2 - m\bar{x}^2\right) = \left(\frac{1}{6m^2}\,(m-1)m(2m-1) - m\bar{x}^2\right)$$

whence $s^2 = \dfrac{1}{12}\left(1 + \dfrac{1}{m}\right)$.

3.14 If $A \times U0 + B = (k \times 1000) + r$

where $\quad 0 \le r < 1000,\ U1 = k + r/1000 + \varepsilon$

where $\varepsilon \ll 0.001$ is the round-off error, and $INT(U1) = r/1000 + \varepsilon$,

$$Y = (U1 - INT(U1)) \times 1000 = r + 1000\varepsilon$$

We require r, therefore set $U1 = INT(r + 1000\varepsilon + \theta)$
where θ is such that $\theta + 1000\varepsilon > 0$, but $\theta + 1000\varepsilon < 1$.
$\theta = 0.5$ will do.
An example is:

1 168 595 82 429 435.964 874.781 . . .

With the additional line we get:

1 168 595 82 429 436 903 . . .

3.15 $x_{i+1} = ax_i + b \pmod{m}$

i.e. $x_{i+1} = ax_i + b - \kappa m \qquad$ for some κ and $0 \le ax_i + b < m$

and so
$$\left(\frac{x_{i+1}}{m}\right) = a\left(\frac{x_i}{m}\right) + \frac{b}{m} - \kappa$$

i.e. $u_{i+1} = a\,u_i + \dfrac{b}{m} \pmod{1}$ and $0 \le a\,u_i + \dfrac{b}{m} < 1$.

3.16 For any $n \ge 0$, $x_{n+1} = ax_n + b \pmod{m}$

Therefore $ax_n + b = \gamma m + x_{n+1}$ for some integral $\gamma \ge 0$ and $0 \le x_{n+1} < m$

and so $x_{n+2} = ax_{n+1} + b \pmod{m}$

$$= a^2x_n + ab - \gamma am + b$$ ~~(mod m)~~

$$= a^2x_n + (a+1)b \pmod{m}$$

and this is the approach which may be used to prove this result by induction on k for any n:

$$x_{n+k+1} = \left(a^{k+1}x_n + \left(\frac{a^k - 1}{a-1}\right)b\,a + b\right) \quad \bmod(m)$$

$$= a^{k+1}x_n + (a^{k+1} - 1)b\,(a-1)^{-1} \quad \bmod(m)$$

Thus every kth term of the original series is another congruential series, with multiplier a^k and increment $(a^k - 1)\,(a-1)^{-1}b$, or, equivalently, $a^k \pmod{m}$ and $(a^k - 1)\,(a-1)^{-1}b \pmod{m}$, respectively.

3.17 The illustration below is taken from Wichmann and Hill (1982a):

Value of U_2 to 1 d.p.	Value of U_1 to 1 d.p.									
	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.0	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.0
0.2	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.0	0.1
0.3	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.0	0.1	0.2
0.4	0.4	0.5	0.6	0.7	0.8	0.9	0.0	0.1	0.2	0.3
0.5	0.5	0.6	0.7	0.8	0.9	0.0	0.1	0.2	0.3	0.4
0.6	0.6	0.7	0.8	0.9	0.0	0.1	0.2	0.3	0.4	0.5
0.7	0.7	0.8	0.9	0.0	0.1	0.2	0.3	0.4	0.5	0.6
0.8	0.8	0.9	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7
0.9	0.9	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8

The values above give the fractional part of $(U_1 + U_2)$. If U_1 and U_2 are independent, then whatever the value of U_2, if U_1 is uniform then so is the fractional part of $(U_1 + U_2)$, and U_2 need not be uniform.

3.20 Note the generalization of this result.

3.21 (a) Numbers in the two half-periods differ by 16:

0 13 2 31 4 17 6 3 8 21 10 7 12 25 14 11
16 29 18 15 20 1 22 19 24 5 26 23 28 9 30 27

We have:

u_r	$9u_r + 13$	u_{r+1}	Binary form of u_{r+1}	Decimal form of $u_{r+1}/32$
0	13	13	01101	0.40625
13	130	2	00010	0.06250
2	31	31	11111	0.96875
31	292	4	00100	0.12500
4	49	17	10001	0.53125
17	166	6	00110	0.18750
6	67	3	00011	0.09375
3	40	8	01000	0.25000
8	85	21	10101	0.65625
21	202	10	01010	0.31250
10	103	7	00111	0.21875
7	76	12	01100	0.37500
12	121	25	11001	0.78125
25	238	14	01110	0.43750
14	139	11	01011	0.34375
11	112	16	10000	0.50000
16	157	29	11101	0.90625
29	274	18	10010	0.56250
⋮	⋮	⋮	⋮	⋮

revealing further clear patterns.

3.22 (a) Procedures are equivalent if the indicator digit is from a generator of period g.

(b) The sequence is: 1, 8, 11, 10, 5, 12, 15, 14, 9, 0, 3, 2, 13, 4, 7, 6, 1. Suppose $g = 4$, for illustration. We start with (1, 8, 11, 10). Let X denote the next number in the sequence, and suppose that if

$$(*)\begin{cases} 0 \le X \le 3 & \text{we replace the 1st stored term} \\ 4 \le X \le 7 & \text{we replace the 2nd stored term} \\ 8 \le X \le 11 & \text{we replace the 3rd stored term} \\ 12 \le X \le 15 & \text{we replace the 4th stored term} \end{cases}$$

This gives:

$$\left.\begin{array}{ll} (1, 8, 11, 10) & \\ (1, 5, 11, 10) & \text{use } 8 \\ (1, 5, 11, 12) & \text{use } 10 \\ (1, 5, 11, 15) & \text{use } 15 \end{array}\right\}$$
etc.

Once the store contains (1, 5, 9, 14) in this example then it enters a (full) cycle. Note that before a cycle can commence, the numbers in the store must correspond, in order, to the four different ranges in (*), and the sequence of numbers for any part of the store must correspond to that in the original sequence for the numbers in that range.

3.23 Omit trailing decimal places, when the x_i are divided by m, to give: $u_i = x_i/m$.

3.25 For suitable α, β, θ,

$$\begin{aligned} x_{i+1} &= (2^{16}+3)x_i + \alpha 2^{31} \\ &= 6x_i + x_i(2^{16}-3) + \alpha 2^{31} \\ &= 6x_i + (2^{16}+3)(2^{16}-3)x_{i-1} + (2^{16}-3)\beta 2^{31} + \alpha 2^{31} \\ &= 6x_i + (2^{32}-9)x_{i-1} + \theta 2^{31} \\ &= 6x_i - 9x_{i-1} + \theta 2^{31} \\ &= 6x_i - 9x_{i-1} \pmod{2^{31}}. \end{aligned}$$

3.26 (b) One-sixth of the time.

$0 \le x_i < m$

therefore $x_n + x_{n-1} = x_{n+1} + \kappa m$ where $\kappa = 0$ or $\kappa = 1$.

Now suppose $x_{n-1} < x_{n+1} < x_n$ (*)

then $x_{n-1} + \kappa m < x_n + x_{n-1} < x_n + \kappa m$

If $\kappa = 1$, this implies $x_n > m$

If $\kappa = 0$, this implies $x_{n-1} < 0$

} neither can occur, so (*) is false.

3.27 (b) No zero values with a multiplicative congruential generator.

3.28 FORTRAN programs are provided by Law and Kelton (1982, p. 227); see also Nance and Overstreet (1975). Some versions of FORTRAN allow one to declare extra-large integers.

3.29 Set $y_n = \theta^n$ and solve the resulting quadratic equation in θ (roots θ_1 and θ_2). The general solution is then $y_n = A\theta_1^n + B\theta_2^n$, where A and B are determined by the values of y_0 and y_1.

3.32 From considering the outcome of the first two tosses,

$$p_n = \tfrac{1}{2}p_{n-1} + \tfrac{1}{4}p_{n-2} \qquad \text{for } n \geq 2.$$

$$np_n = \frac{n}{2}p_{n-1} + \frac{n}{4}p_{n-2}$$
$$= \frac{(n-1)}{2}p_{n-1} + \tfrac{1}{2}p_{n-1} + \frac{(n-2)}{4}p_{n-2} + \tfrac{1}{2}p_{n-2} \qquad \text{for } n \geq 2$$

Summing over n: if $\mu = \sum_{n=1}^{\infty} np_n$,

$$\mu - 2p_2 - p_1 = \frac{\mu}{2} + \frac{\mu}{4} + \tfrac{1}{2}(1-p_1) + \tfrac{1}{2};\ p_1 = 0,\ p_2 = \tfrac{1}{4},\ \mu = 6.$$

Chapter 4

4.2

```
10   LET P = .5
20   LET Q = 1-P
30   LET S = 0
40   RANDOMIZE
50   LET U = RND
60   FOR I = 1 TO 3
70    IF U < P THEN 100
80     LET U = (1-U)/Q
90    GOTO 120
100    LET U = U/P
110    LET S = S+1
120   NEXT I
130   PRINT S
140   END
```

4.6 $N_1 = (-2\log_e U_1)^{1/2}\cos 2\pi U_2$

$N_2 = (-2\log_e U_1)^{1/2}\sin 2\pi U_2$

$\dfrac{n_2}{n_1} = \tan(2\pi u_2);\ u_1 = \exp[-\tfrac{1}{2}(n_1^2+n_2^2)]$

$$\begin{vmatrix} \dfrac{\partial u_1}{\partial n_1} & \dfrac{\partial u_1}{\partial n_2} \\ \dfrac{\partial u_2}{\partial n_1} & \dfrac{\partial u_2}{\partial n_2} \end{vmatrix} = \begin{vmatrix} -n_1u_1 & -n_2u_1 \\ \dfrac{-n_2}{n_1^2 2\pi\sec^2(2\pi u_2)} & \dfrac{1}{2\pi n_1\sec^2(2\pi u_2)} \end{vmatrix}$$

$$= \frac{u_1(1+n_2^2/n_1^2)}{2\pi\sec^2(2\pi u_2)} = \frac{u_1}{2\pi} = \frac{\exp[-\tfrac{1}{2}(n_1^2+n_2^2)]}{2\pi}, \text{ as required.}$$

4.8 The simplest case is when $n = 2$, resulting in the triangular distribution:

$$\begin{aligned} f(x) &= x \quad &\text{for} \quad 0 \leq x \leq 1 \\ f(x) &= 2 - x \quad &\text{for} \quad 1 \leq x \leq 2 \end{aligned}$$

For illustrations and applications, see Section 5.5.

4.9 Define $Y_0 = 0$.
$K = i$ if and only if $Y_i \leq 1 < Y_{i+1}$ for $i \geq 0$.
Thus $K \leq k$ if and only if $Y_{k+1} > 1$. Therefore

$$\Pr(K \leq k) = \Pr(Y_{k+1} > 1) = \int_1^\infty \frac{\lambda(\lambda y)^k e^{-\lambda y} \mathrm{d}y}{k!}$$

and $\Pr(K = k) = \Pr(K \leq k) - \Pr(K \leq k-1)$ for $k \geq 0$

$[\Pr(K \leq -1) \equiv \Pr(Y_0 > 1) = 0]$

$$= \Pr(Y_{k+1} > 1) - \Pr(Y_k > 1)$$

$$= \int_1^\infty \left\{ \frac{\lambda^{k+1} y^k e^{-\lambda y}}{k!} - \frac{\lambda^k y^{k-1} e^{-\lambda y}}{(k-1)!} \right\} \mathrm{d}y$$

But

$$\int_1^\infty \frac{\lambda^{k+1} y^k e^{-\lambda y}}{k!} \mathrm{d}y = -\frac{\lambda^k}{k!} \int_1^\infty y^k \mathrm{d}(e^{-\lambda y})$$

$$= -\left[\frac{\lambda^k}{k!} y^k e^{-\lambda y} \right]_1^\infty + \frac{\lambda^k}{(k-1)!} \int_1^\infty e^{-\lambda y} y^{k-1} \mathrm{d}y$$

and so (integrals are finite)

$$\Pr(K = k) = \frac{e^{-\lambda} \lambda^k}{k!} \qquad \text{for } k \geq 0.$$

4.10 Lenden–Hitchcock (1980) considered the following four-dimensional method.

Let N_i, $1 \leq i \leq 4$, be independent random variables, each with the half-normal density:

$$f_{N_i}(x) = \sqrt{\left(\frac{2}{\pi}\right)} e^{-x^2/2} \qquad \text{for } x \geq 0.$$

Changing from Cartesian to polar co-ordinates we get:

$$(*) \begin{cases} N_1 = R \cos \Theta_1 \cos \Theta_2 \cos \Theta_3 \\ N_2 = R \sin \Theta_1 \cos \Theta_2 \cos \Theta_3 & 0 < \Theta_i < \dfrac{\pi}{2}, \text{ for } 1 \leq i \leq 4 \\ N_3 = R \sin \Theta_2 \cos \Theta_3 & 0 < R < \infty. \\ N_4 = R \sin \Theta_3 \end{cases}$$

The Jacobian of the transformation is $R^3 \cos\Theta_2 \cos^2\Theta_3$, and so

$$f_{R,\Theta_1,\Theta_2,\Theta_3}(r, \theta_1, \theta_2, \theta_3) = \frac{4}{\pi^2} r^3 e^{-r^2/2} \cos\theta_1 \cos^2\theta_2.$$

Thus R, Θ_1, Θ_2 and Θ_3 are independent, and to obtain the $\{N_i\}$ we simply simulate R, Θ_1, Θ_2 and Θ_3 using their marginal distributions and transform back to the $\{N_i\}$ by (*). R^2 has a χ^2_4 distribution, and so is readily simulated, as described in Section 4.3.

(Had a three-dimensional generalization been used, R^2 would have had a χ^2_3 distribution, which is less readily simulated.) Lenden–Hitchcock suggested that the four-dimensional method is more efficient than the standard two-dimensional one.

4.12 Let $\Psi = \Pr(\text{Accept } (V_2, V_3) \text{ and Accept } (V_1, V_2))$

$$= \Pr(V_2^2 + V_3^2 \le 1 \text{ and } V_2^2 + V_1^2 \le 1)$$

$$f_{V^2}(x) = \frac{1}{2x^{1/2}} \quad \text{for } 0 \le x \le 1; \qquad F_{V^2}(x) = x^{1/2}$$

Therefore $$\Psi = \int_0^1 \Pr(V_3^2 \le 1 - x) \Pr(V_1^2 \le 1 - x) \frac{1}{2x^{1/2}} dx$$

$$= \frac{1}{2}\int_0^1 (x^{-1/2} - x^{1/2}) dx = 1 - \frac{1}{3} = \frac{2}{3} \ne \frac{\pi^2}{16}.$$

4.13 Logistic.

4.14 It is more usual to obtain the desired factorization using Choleski's method (see Conte and de Boor, 1972, p. 142), which states that we can obtain the desired factorization with a lower triangular matrix **A**. The elements of **A** can be computed recursively as follows.

Let $\mathbf{\Sigma} = \{\sigma_{ij}\}$, $\mathbf{A} = \{a_{ij}\}$. We have $a_{ij} = 0$ for $j > i$, and so $\sigma_{ij} = \sum_{k=1}^{\min(i,j)} a_{ik} a_{jk}; \sigma_{11} = a_{11}^2$, so that $a_{11} = \sigma_{11}^{1/2}$, and then $a_{i1} = (\sigma_{i1}/\sigma_{11}^{1/2})$ for $i > 1$. This gives the first column of **A**, which may be used to give the second column, and so on. Given the first $(j-1)$ columns of **A**,

$$a_{jj} = \left(\sigma_{jj} - \sum_{k=1}^{j-1} a_{jk}^2\right)^{1/2}$$

and $$a_{ij} = \left(\sigma_{ij} - \sum_{k=1}^{j-1} a_{ik} a_{jk}\right) / a_{jj} \qquad \text{for } i > j.$$

4.16 The result is easily shown by the transformation-of-variable theory, with the Jacobian of the transformation $= (1 - \rho^2)^{1/2}$. Alternatively, small-scale study resulting from data collected from a provincial French practice, see Example 8.1.

1.7 If a model did not result in a simplification then it would not be a model. Clearly models such as that of Exercise 1.6 ignore features which would make the models more realistic. Thus one might expect more women patients than men (adults) in the morning, as compared with the afternoon. Were this true, and if there was a sex difference regarding consultation times, or lateness factors, the model should be modified accordingly.

1.9 A system with small mean waiting time may frequently give rise to very small waiting times, but occasionally result in very large waiting times. An alternative system with slightly larger mean waiting time, but no very large waiting times, could be preferable. An example of this kind is discussed by Gross and Harris (1974, p. 430). In some cases a multivariate response may be of interest (see Schruben, 1981).

1.10 The following two examples are taken from Shannon and Weaver (1964, pp. 43–44).

(a) Words chosen independently but with their appropriate frequencies: 'Representing and speedily is an good apt or come can different natural here he the a in came the to of to expert gray came to furnishes the line message had be these.'

(b) If we simulate to match the first-order transition frequencies, i.e. matching the frequencies of what follows what, we get: 'The head and in frontal attack on an English writer that the character of this point is therefore another method for the letters that the time of who ever told the problem for an unexpected.'

Chapter 2

2.1 $X = -\log_e U; f_X(x) = f_Y(y)\left|\frac{\mathrm{d}y}{\mathrm{d}x}\right|$

in general, and so here,

$$f_X(x) = 1\left|\frac{\mathrm{d}u}{\mathrm{d}x}\right| = \frac{1}{u^{-1}}$$

but $u = e^{-x}$, and so $f_X(x) = e^{-x}$, for $x \geq 0$.

2.2 Note that for the exponential, gamma and normal cases X remains, respectively, exponential, gamma and normal.

2.3 $f_X(x) = e^{-x} \qquad \text{for } x \geq 0$

$W = \gamma X^{1/\beta}$

the joint p.d.f. of Y_1 and X_1 can be written as the product of the conditional p.d.f. of $Y_1|X_1$ and the marginal distribution of X_1, to yield, directly:

$$f_{Y_1X_1}(y, x) = \frac{\exp[-\frac{1}{2}(y-\rho x)^2/(1-\rho^2)]}{(1-\rho^2)^{1/2}\sqrt{(2\pi)}} \frac{\exp[-\frac{1}{2}x^2]}{\sqrt{(2\pi)}}$$

(See also the solution to Exercise 6.5.)

4.17 Median $= m$ if and only if one value $= m$, $(n-1)$ of the other values are less than m, and $(n-1)$ are greater than m. The value to be m can be chosen $(2n-1)$ ways; the values to be $< m$ can be chosen $\binom{2n-2}{n-1}$ ways, hence,

$$f_M(m)\,\mathrm{d}m = (2n-1)\binom{2n-2}{n-1} m^{n-1}(1-m)^{n-1}\,\mathrm{d}m$$

i.e. M has a $B_e(n, n)$ distribution.

4.19 $$\Pr(X = k) = \frac{1}{k!}\int_0^\infty \frac{e^{-\lambda}\lambda^k e^{-\theta\lambda}\theta^n\lambda^{n-1}}{\Gamma(n)}\,\mathrm{d}\lambda$$

$$= \frac{\theta^n}{k!\,\Gamma(n)}\int_0^\infty \lambda^{n+k-1}e^{-\lambda(\theta+1)}\,\mathrm{d}\lambda$$

$$= \frac{\theta^n}{k!\,\Gamma(n)}\frac{\Gamma(n+k)}{(\theta+1)^{n+k}} \qquad \text{as required.}$$

4.21 $y = e^x$; $\dfrac{\mathrm{d}y}{\mathrm{d}x} = e^x = y$; $x = \log_e(y)$

$$f_Y(y) = \frac{1}{y\sigma\sqrt{(2\pi)}}\exp\left\{-\frac{1}{2}\left(\frac{\log_e y - \mu}{\sigma}\right)^2\right\} \qquad \text{for } y \geq 0.$$

4.22 $$M(\theta) = \sum_{k=1}^{\infty} e^{\theta k}p_k = -\frac{1}{\log(1-\alpha)}\sum_{k=1}^{\infty}\frac{(\alpha e^\theta)^k}{k}$$

Note that $$\mathrm{d}M(\theta)/\mathrm{d}\theta = -\frac{1}{\log(1-\alpha)}\sum_{j=0}^{\infty}(\alpha e^\theta)^j\alpha e^\theta$$

i.e. $$\frac{\mathrm{d}M(\theta)}{\mathrm{d}\theta} = \frac{\alpha e^\theta}{\log(1-\alpha)(\alpha e^\theta - 1)} \qquad \text{for } \alpha e^\theta < 1$$

Therefore $$M(\theta) = \kappa + \frac{1}{\log(1-\alpha)}\log(1-\alpha e^\theta)$$

$$M(0) = 1 \text{ so } \kappa = 0.$$

$$\Pr(X = k) = -\frac{1}{\log(1-\alpha)}\int_0^\alpha y^{k-1}\,\mathrm{d}y = -\frac{\alpha^k}{k\log(1-\alpha)} \qquad \text{for } k \geq 1.$$

This question is continued in Exercise 5.14.

Chapter 5

5.1 Let $W = 1 - U$; $f_W(w) = f_U(u)\left|\frac{du}{dw}\right|$

$$\left|\frac{du}{dw}\right| = 1 \qquad \text{and the result is proved.}$$

5.2
$$F_{\tilde{X}}(w) = \Pr(\tilde{X} \le w) = \Pr(\tilde{X} \le w | \tilde{X} = X)\Pr(\tilde{X} = X) + \Pr(\tilde{X} \le w | \tilde{X} = -X)\Pr(\tilde{X} = -X)$$
$$= \tfrac{1}{2}\Pr(X \le w) + \tfrac{1}{2}\Pr(-X \le w)$$
$$= \tfrac{1}{2}\Pr(X \le w) + \tfrac{1}{2}\Pr(X \ge -w)$$

Therefore if $w \ge 0$,

$$F_{\tilde{X}}(w) = \frac{1}{2}\int_0^w f_X(x)\,dx + \frac{1}{2} = \Phi(w)$$

and if $w \le 0$,

$$F_{\tilde{X}}(w) = 0 + \tfrac{1}{2}\Pr(X \ge -w) = \Phi(-w).$$

5.3 Poisson random variables with large mean values, μ, say, can be simulated as the sum of independent Poisson variables with means which sum to μ.

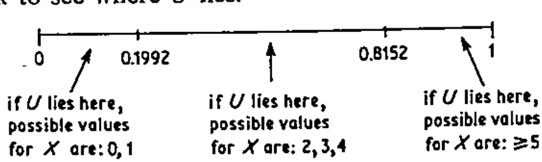
5.4 We can take a Poisson random variable X, with mean 3 as an illustration:

i	0	1	2	3	4	5	6
$\Pr(X = i)$	0.0498	0.1494	0.2240	0.2240	0.1680	0.1008	0.0504

i	7	8	9	10	≥ 11
$\Pr(X = i)$	0.0216	0.008	0.003	0.0008	0.0002

(b) Here we can set $\theta = 2$, say; $\Pr(X \le 2) = 0.4232$.
As a first stage, check to see if $U > 0.4232$. If so, then it is not necessary to check U against $\Sigma_{i=0}^{j}\Pr(X = i)$, for $j \le \theta = 2$.

(c) Here we could take $\theta_1 = 1$ and $\theta_2 = 4$
$\Pr(X \le 1) = 0.1992$; $\Pr(X \le 4) = 0.8152$, and as a first stage we check to see where U lies:



(d) When the probabilities are ordered we have:

i	2	3	4	1	5	6
$\Pr(X = i)$	0.224	0.224	0.168	0.1494	0.1008	0.0504

i	0	7	8	9	10	≥ 11
$\Pr(X = i)$	0.0498	0.0216	0.008	0.003	0.0008	0.0002

Ordering is a time-consuming operation but once it is done we obtain the obvious benefit of checking the most likely intervals first. As a first step one might check to see if $U \leq 0.616$. If so, X takes one of the values 2, 3 or 4.

5.5 (a) Illustrate $F(x)$ by means of a diagram. Note the symmetry. If $U \leq 0.5$, set $X = \sqrt{(2U)}$; if $U \geq 0.5$, set $X = 2 - \sqrt{(2-2U)}$.
(b) We also obtain such an X by setting $X = U_1 + U_2$, where U_1 and U_2 are independent, identically distributed $U(0, 1)$ random variables (see Exercise 4.8).

5.7 $$F_X(x) = 6\int_0^x y(1-y)\,dy = 6\left[\frac{y^2}{2} - \frac{y^3}{3}\right]_0^x = 3x^2 - 2x^3$$

Therefore set $U = 3X^2 - 2X^3$ and solve this cubic equation in X.

5.8 (a) $F(x) = \dfrac{1}{\pi}\displaystyle\int_{-\infty}^{x} \frac{dy}{(1+y^2)}$, $-\infty \leq x \leq \infty$.

Set $y = \tan\theta$; $F(x) = \dfrac{1}{\pi}\displaystyle\int_{-\pi/2}^{\tan^{-1}x} d\theta = \dfrac{1}{\pi}\left(\tan^{-1}x + \dfrac{\pi}{2}\right)$

for $-\dfrac{\pi}{2} \leq \theta \leq \dfrac{\pi}{2}$

Therefore set $U\pi = \tan^{-1}x + \dfrac{\pi}{2}$,

i.e. $\tan(U\pi + \pi/2) = X$
or, equivalently,
$X = \tan(\pi U)$

(b) We saw, from Section 4.2.1, that we can write
$N_1 = (-2\log_e U_1)^{1/2} \sin 2\pi U_2$
$N_2 = (-2\log_e U_1)^{1/2} \cos 2\pi U_2$
where U_1 and U_2 are independent $U(0, 1)$ random variables. Thus $N_1/N_2 = \tan(2\pi U_2)$, which clearly has the same distribution as $X = \tan(\pi U)$ above, i.e. the standard Cauchy.

5.9 Let $u^{-1} = 1 + \exp(-2a_1\tilde{x}(1 + a_2\tilde{x}^2))$
Therefore we need to solve

$$\log_e(u^{-1} - 1) = -2a_1\tilde{x}(1 + a_2\tilde{x}^2)$$

Use the result: If $ax^3 + x - b = 0$, $x = c - 1/(3ac)$
where

$$2ac^3 = b + \left(b^2 + \frac{4}{27a}\right)^{1/2} \qquad \text{(see Page, 1977).}$$

5.10 See Kemp and Loukas (1978a, b). They suggest truncating the distribution—e.g. to give values of $\{p_{ij}\}$ summing, over i and j, to give unity to, say, 4 decimal places. Then use the approach of Exercise 5.4(d). They found this approach faster than first simulating a marginal variable and then simulating the bivariate distribution via the conditional distribution.

5.11 One approach is via the conditional and marginal distributions, as above, in the comments on Exercise 5.10.

5.12 $\Pr(Y = i-1) = e^{-\lambda(i-1)} - e^{-\lambda i} = e^{-\lambda i}(e^{\lambda} - 1) \qquad \text{for } i \geq 1$

$$= e^{-\lambda(i-1)}(1 - e^{-\lambda}) \qquad \text{for } i \geq 1.$$

Hence, from Section 2.6, $(Y+1)$ has a geometric distribution with $q = e^{-\lambda}$. This relationship is not surprising since both geometric and exponential random variables measure waiting-times, as stated in Section 2.10.

5.13 (a) $F(x) = 1/(1 + e^{-x})$
Therefore set $U = 1/(1 + e^{-X})$

$$U^{-1} = 1 + e^{-X}$$

$$X = -\log_e(U^{-1} - 1).$$

(b) $$F(x) = \frac{\beta}{\gamma^\beta}\int_0^x w^{\beta-1}\exp[-(w/\gamma)^\beta]\,dw$$

$$= \Big[\exp[-(w/\gamma)^\beta]\Big]_x^0 = 1 - \exp[-(x/\gamma)^\beta]$$

Therefore set $U = 1 - \exp[-(X/\gamma)^\beta]$

i.e. $\exp[-(X/\gamma)^\beta] = 1 - U$; $\quad -\left(\frac{X}{\gamma}\right)^\beta = \log_e(1 - U)$,

$$X = \gamma(-\log_e(1 - U))^{1/\beta}$$

or equivalently, and more simply (see Exercise 5.1)

$$X = \gamma(-\log_e U)^{1/\beta}$$

(c) Set $U = 1-(k/X)^a$; $1-U = (k/X)^a$

$$k(1-U)^{-1/a} = X$$

or equivalently, as in (b)

$$X = kU^{-1/a}$$

(d) Set $U = \exp(-\exp((\xi - X)/\theta)$

$$-\log_e U = \exp((\xi - X)/\theta)$$
$$\theta \log_e(-\log_e U) = \xi - X,$$
$$X = \xi - \theta \log_e(-\log_e U)$$

5.14 To complete the simulation of a random variable with the logarithmic distribution of Exercise 4.22, we need to simulate the random variable Y, for which

$$F_Y(y) = \log(1-y)/\log(1-\alpha) \qquad \text{for } 0 \leq y \leq \alpha.$$

To do this by the inversion method, set $U = \log(1-Y)/\log(1-\alpha)$, i.e. $\log(1-Y) = U\log(1-\alpha)$; $1-Y = (1-\alpha)^U$;

$$Y = 1-(1-\alpha)^U$$

5.15 (a) First we find $f(x, y)$:

$$f(x, y) = f(x)f(y)(1-\alpha(1-F(x))(1-F(y))) + f(x)F(y)\alpha f(y) \times (1-F(x))$$
$$= \alpha F(x)f(y)f(x)(1-F(y)) - \alpha f(x)f(y)F(x)F(y)$$

Now make the transformation of variable: $U = F(X)$, $V = F(Y)$:

$$f(u, v) = 1-\alpha(1-u)(1-v) + \alpha v(1-u) + \alpha u(1-v) - \alpha uv$$
$$= 1-\alpha(1-2u)(1-2v)$$

(b) $f(u, v) = (1-\alpha \log u)(1-\alpha \log v) - \alpha) v u^{-\alpha \log v}$

(c) $f(u, v) = \dfrac{\pi}{2} \dfrac{\{1+\tan^{-2}(\pi u)\}\{1+\tan^{-2}(\pi v)\}}{\{1+\tan^{-2}(\pi u)+\tan^{-2}(\pi v)\}^{3/2}}$

5.16 Rejection, with a high probability of rejection, can be applied simply by generating a uniform distribution of points over an enveloping rectangle, and accepting the abscissae of the points that lie below the curve. High rejection here is not too important as this density is only sampled with probability 0.0228.

5.17 $X = \cos \pi U$; $f_X(x) = \left|\dfrac{du}{dx}\right|$

$$\frac{dx}{du} = -\pi \sin(\pi u); \quad f_X(x) = \frac{1}{\pi \sin \pi u} = \frac{1}{\pi\sqrt{(1-x^2)}}$$

for $-1 \leq x \leq 1$.

To simulate X without using the 'cos' function we can use the same approach as in the Polar Marsaglia method:

$$\cos \pi U = 2\cos^2\left(\frac{\pi U}{2}\right) - 1$$

Thus, if U_1, U_2 are independent $U(0, 1)$ random variables

$$\text{if } U_1^2 + U_2^2 < 1, \text{ set } X = \frac{2U_1^2}{U_1^2 + U_2^2} - 1 = \left(\frac{U_1^2 - U_2^2}{U_1^2 + U_2^2}\right).$$

5.18 Using a rectangular envelope gives an acceptance probability of 2/3. Using a symmetric triangular envelope gives an acceptance probability of 2/3 also, but the acceptance probability of 8/9 results from using a symmetric trapezium. Simulation from a trapezium density is easily done using inversion.

5.19 For the exponential envelope we simulate from the half-logistic density:

$$f(x) = 2e^{-x}/(1+e^{-x})^2 \qquad \text{for } x \geq 0$$

We need to choose $k > 1$ so that $ke^{-x} > 2e^{-x}/(1+e^{-x})^2$ for all $x \geq 0$, i.e. choose k so that $k > 2(1+e^{-x})^{-2}$ for all $x \geq 0$, and this is done for smallest k by setting $k = 2$ $(x = \infty)$.
If U_1, U_2 are independent $U(0, 1)$ random variables, set $X = -\log_e U_1$ if $U_2 < (1+U_1)^{-2}$. Finally transform to $(-\infty, \infty)$ range.

5.20
$$\left.\begin{aligned} f(x) &= \frac{1}{\sqrt{(2\pi)}} e^{-x^2/2} \\ h(x) &= e^{-x}(1+e^{-x})^{-2} \end{aligned}\right\} \qquad -\infty \leq x \leq \infty$$

Set $q(x) = \dfrac{1}{\sqrt{(2\pi)}} e^{-x^2/2}(1+e^{-x})^2 e^x$.

If $l(x) = \log_e q(x)$

$$l(x) = -\frac{1}{2}\log_e(2\pi) + x - \frac{x^2}{2} + 2\log_e(1+e^{-x})$$

$$\frac{dl(x)}{dx} = 1 - x - \frac{2}{(1+e^x)} = 0 \qquad \text{when } x = 0$$

$$\frac{d^2l(x)}{dx^2} = -1 + \frac{2e^x}{(1+e^x)^2} < 0 \qquad \text{when } x = 0$$

Therefore $x = 0$ maximizes $q(x)$, and $k = \max_x (q(x)) = \dfrac{4}{\sqrt{(2\pi)}}$ $= 1.596$.

To operate the rejection method here we need to simulate from $h(x)$, and this is readily done by the inversion method of Exercise 5.13(a); probability of rejection $= 1 - 1/k = 0.37$. This may seem surprisingly high, but recall that the two distributions have different variances (1 and $\pi^2/3$) and a better approach would be to use a logistic distribution of unit variance (see Fig. 5.14).

5.21 $k = \max_x \left\{ \sqrt{\left(\frac{2}{\pi}\right)} \exp[-x^2/2]/(\lambda \exp[-\lambda x]) \right\}$

i.e. $k = \max_x \left\{ \left(\frac{1}{\lambda} \sqrt{\left(\frac{2}{\pi}\right)} \exp(\lambda x - x^2/2) \right\} $

Let $y = \lambda x - x^2/2$; $\frac{dy}{dx} = \lambda - x = 0$ when $x = \lambda$, $\frac{d^2y}{dx^2} = -1$.

Hence k is obtained for $x = \lambda$, to give:

$$k = \frac{1}{\lambda} \sqrt{\left(\frac{2}{\pi}\right)} \exp(\lambda^2/2)$$

The method becomes: accept $X = -\frac{1}{\lambda} \log_e U_1$ if

$U_1 U_2 \leq \exp[-(\lambda^2 + X^2)/2]$, for independent $U(0, 1)$ random variables U_1 and U_2.

Probability of rejection $= 1 - \lambda \sqrt{\left(\frac{\pi}{2}\right)} \exp(-\lambda^2/2)$

Let $y = \lambda \exp(-\lambda^2/2)$

let $z = \log y = \log \lambda - \lambda^2/2$; $\frac{dz}{d\lambda} = \frac{1}{\lambda} - \lambda = 0$ when $\lambda = 1$.

$\frac{d^2z}{d\lambda^2} = -\frac{1}{\lambda^2} - 1, < 0$ when $\lambda = 1$

Thus taking $\lambda = 1$ minimizes the probability of rejection.

5.22 (a) Use the inversion method:

$$F(x) = \lambda\mu \int_0^x y^{\lambda-1} (\mu + y^\lambda)^{-2} \, dy = \Big[\mu(\mu + y^\lambda)^{-1} \Big]_x^0$$

$$= 1 - \mu(\mu + x^\lambda)^{-1} \qquad \text{for } x \geq 0.$$

Set $U = 1 - \mu(\mu + X^\lambda)^{-1}$; $1 - U = \mu(\mu + X^\lambda)^{-1}$

or equivalently

$$\frac{\mu}{U} = \mu + X^\lambda; \quad X = \left\{ \mu \left(\frac{1}{U} - 1 \right) \right\}^{1/\lambda}$$

$(W/\gamma)^\beta = X$

$f_W(w) = f_X(x)|\mathrm{d}x/\mathrm{d}w|$

$\mathrm{d}w/\mathrm{d}x = \dfrac{\gamma}{\beta}\, x^{1/\beta - 1}$

$f_W(w) = \dfrac{\beta e^{-x}}{\gamma}\, x^{1-1/\beta} = \dfrac{\beta}{\gamma}\, e^{-(w/\gamma)^\beta} \left(\dfrac{w}{\gamma}\right)^{(\beta-1)}$

i.e. $f_W(w) = \dfrac{\beta}{\gamma^\beta}\, w^{\beta-1} e^{-(w/\gamma)^\beta}$ for $w \geq 0$.

2.4 $Y = N^2$; $\Pr(0 \leq Y \leq y) = \Pr(-y^{1/2} \leq N \leq y^{1/2})$

$$= \Phi(y^{1/2}) - \Phi(-y^{1/2})$$

and so

$$f_Y(y) = \tfrac{1}{2}y^{-1/2}\phi(y^{1/2}) + \tfrac{1}{2}y^{-1/2}\phi(-y^{1/2}) = y^{-1/2}\phi(y^{1/2})$$
$$= y^{-1/2}e^{-y/2}/\sqrt{(2\pi)} \qquad \text{for } y \geq 0, \text{ i.e. } \chi_1^2.$$

Note that $Y = N^2$ is not 1–1 and so rote application of Equation (2.3) gives the wrong answer.

2.5 $M_Y(\theta) = \mathscr{E}[e^{Y\theta}] = \displaystyle\int_0^\infty \frac{y^{-1/2}e^{y(\theta-1/2)}}{\sqrt{(2\pi)}}\,\mathrm{d}y$

Suppose that $\theta < \frac{1}{2}$. Let $z = -2y(\theta - \frac{1}{2})$,
i.e. $z = y(1-2\theta)$; $\mathrm{d}z = \mathrm{d}y(1-2\theta)$,

$$M_Y(\theta) = \int_0^\infty \left(\frac{z}{1-2\theta}\right)^{-1/2} e^{-z/2} \frac{\mathrm{d}z}{\sqrt{(2\pi)}(1-2\theta)}$$
$$= \frac{(1-2\theta)^{-1/2}}{\sqrt{(2\pi)}} \int_0^\infty z^{-1/2}e^{-z/2}\,\mathrm{d}z,$$

i.e. $M_Y(\theta) = (1-2\theta)^{-1/2}$ (see also Table 2.1).

Hence, in the notation of this question,

$M_{\sum_{i=1}^{n} Y_i^2}(\theta) = (1-2\theta)^{-n/2}$

i.e. (from Table 2.1), the m.g.f. of a χ_n^2 random variable. (See also Exercise 2.21.)

2.7 Solution:

If $S = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} Y_i$ then $f_S(s) = \dfrac{1}{\pi(1+s^2)}$ for $-\infty \leq s \leq \infty$, i.e. S has the same Cauchy p.d.f. as the component $\{Y_i\}$ random variables.

2.8 (a) $N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$
(b) Poisson with parameter $(\lambda+\mu)$
(c) If $\lambda=\mu$ the solution is given by Example 2.6, with a generalization given in Exercise 2.6. If $\lambda \neq \mu$,

$$M_{X+Y}(\theta)=\left(\frac{\lambda}{\lambda-\theta}\right)\left(\frac{\mu}{\mu-\theta}\right)=\frac{1}{(\lambda-\theta)}\frac{\lambda\mu}{(\mu-\lambda)}+\frac{1}{(\mu-\theta)}\frac{\lambda\mu}{(\lambda-\mu)}$$

and so,

$$f_{X+Y}(z)=\left(\frac{\lambda\mu}{\mu-\lambda}\right)e^{-\lambda z}+\left(\frac{\lambda\mu}{\lambda-\mu}\right)e^{-\mu z} \qquad \text{for } z \geq 0$$

i.e. a mixture of exponential densities.

While the solutions for (a) and (b) also follow easily from using generating functions, the solution to any one of these parts of this question is also readily obtained from using the convolution integral.

For (a), if the convolution integral is used, note that:

$$\int_{-\infty}^{\infty}\exp\{-\tfrac{1}{2}(ax^2+2bx+c)\}\,dx=\sqrt{\left(\frac{2\pi}{a}\right)}\exp\left(\frac{b^2-ac}{2a}\right).$$

2.9 $$\Pr(X=r|X+Y=n)=\frac{\Pr(X=r \text{ and } Y=n-r)}{\Pr(X+Y=n)}$$

$$=\frac{e^{-\lambda}\lambda^r}{r!}\frac{e^{-\mu}\mu^{n-r}}{(n-r)!}\frac{n!}{e^{-(\lambda+\mu)}(\lambda+\mu)^n}=\binom{n}{r}\frac{\lambda^r\mu^{n-r}}{(\lambda+\mu)^n}$$

i.e. the conditional distribution of X is binomial:

$$B(n, \lambda/(\lambda+\mu)).$$

2.10 $\Pr(Z\leq z)=\Pr(\max(X,Y)\leq z)=\Pr(X\leq z \text{ and } Y\leq z)$

$$F_Z(z)=F_X(z)F_Y(z).$$

2.11 $\Pr(Y\leq y)=(1-e^{-y})^n$

Therefore $f_Y(y)=n(1-e^{-y})^{n-1}e^{-y}$

Now $M_Z(\theta)=\dfrac{1}{(1-\theta)}\dfrac{2}{(2-\theta)}\cdots\dfrac{n}{(n-\theta)}=\sum_{i=1}^{n}\left(\dfrac{i}{i-\theta}\right)\prod_{\substack{j=1\\ j\neq i}}^{n}\left(\dfrac{j}{j-i}\right)$

Note that

$$\prod_{\substack{j=1\\ j\neq i}}^{n}\left(\frac{j}{j-i}\right)=\frac{1\cdot 2\cdot\ldots(i-1)(i+1)(i+2)\ldots n}{(1-i)(2-i)\ldots(-1)1\cdot 2\cdot\ldots(n-i)}=(-1)^{i-1}\binom{n}{i}$$

therefore $M_Z(\theta)=\sum_{i=1}^{n}(-1)^{i-1}\binom{n}{i}\left(\dfrac{i}{i-\theta}\right)$

and so $f_Z(z) = \sum_{i=1}^{n} \binom{n}{i}(-1)^{i-1} i e^{-iy}$

and as $n\binom{n-1}{i-1} = i\binom{n}{i}$ then it is clear that Y and Z have the same distribution.

Note that

$$M_Y(\theta) = n\int_0^\infty (1-e^{-y})^{n-1} e^{-y} e^{\theta y}\,\mathrm{d}y$$

If we let $u = e^{-y}$, $\mathrm{d}u = -e^{-y}\mathrm{d}y$, and so

$$M_Y(\theta) = n\int_0^1 (1-u)^{n-1} u^{-\theta}\,\mathrm{d}u,$$

i.e. is proportional to the beta integral, so that

$$M_Y(\theta) = \frac{n\Gamma(n)\Gamma(1-\theta)}{\Gamma(n-\theta+1)}$$
$$= \frac{n!}{(1-\theta)(2-\theta)\ldots(n-\theta)}$$

as above. To prove this result by induction, note that if

$$W = X/(n-1), f_W(w) = (n+1)e^{-(n+1)w} \qquad \text{for } w \geq 0.$$

Let $Z_{n+1} = Z_n + W$, in an obvious notation, then using the convolution integral

$$f_{Z_{n+1}}(z) = n\int_0^z (1-e^{-y})^{n-1} e^{-y}(n+1)e^{-(n+1)(z-y)}\,\mathrm{d}y$$
$$= n(n+1)e^{-(n+1)z}\int_0^z (1-e^{-y})^{n-1} e^{ny}\,\mathrm{d}y$$
$$= n(n+1)e^{-(n+1)z}\int_0^z (e^{y}-1)^{n-1} e^{y}\,\mathrm{d}y$$
$$= (n+1)e^{-(n+1)z}(e^{z}-1)^n = (n+1)(1-e^{-z})^n e^{-z}.$$

Thus if the result is true for n, it is true for $(n+1)$. But it is clearly true for $n = 0$, and so it is true for $n \geq 0$.

Both Y and Z may be interpreted as the time to extinction for a population of size n, living according to the rules of a linear birth process (see Cox and Miller, 1965, p. 156, or Bailey, 1964, p. 88, and the solution to Exercise 8.4).

2.12 $\left.\begin{array}{l} X_1 = Y_1 - Y_2 \\ X_2 = Y_1 + Y_2 \end{array}\right\}$ $\quad \dfrac{\partial x_1}{\partial y_1} = 1;\ \dfrac{\partial x_1}{\partial y_2} = -1$ $\quad \dfrac{\partial x_2}{\partial y_1} = 1;\ \dfrac{\partial x_2}{\partial y_2} = 1$

$$f_{X_1X_2}(x_1x_2) = f_{Y_1Y_2}(y_1y_2)\begin{vmatrix} \dfrac{\partial y_1}{\partial x_1} & \dfrac{\partial y_1}{\partial x_2} \\ \dfrac{\partial y_2}{\partial x_1} & \dfrac{\partial y_2}{\partial x_2} \end{vmatrix}.$$

$$= \frac{\lambda^2}{2} e^{-\lambda(y_1+y_2)} = \frac{\lambda^2}{2} e^{-\lambda x_2}.$$

It is now that this problem becomes tricky—we must determine the possible region for X_1 and X_2. Clearly, X_1 and X_2 are not independent, and the ranges of each depend upon the value taken by the other.

If $x_1 \geq 0$, then $x_2 \geq x_1$, while if $x_1 \leq 0$, $x_2 \geq -x_1$. So, to obtain the marginal distributions:

$$f_{X_1}(x_1) = \lambda^2 \int_{x_1}^{\infty} \frac{e^{-\lambda x_2}}{2} \mathrm{d}x_2 \qquad \text{for } x_1 \geq 0$$

$$= \lambda^2 \int_{-x_1}^{\infty} \frac{e^{-\lambda x_2}}{2} \mathrm{d}x_2 \qquad \text{for } x_1 \leq 0$$

$$= \frac{\lambda}{2} e^{-\lambda x_1} \qquad \text{for } x_1 \geq 0$$

$$= \frac{\lambda}{2} e^{\lambda x_1} \qquad \text{for } x_1 \leq 0$$

and $$f_{X_2}(x_2) = \int_{-x_2}^{x_2} \frac{\lambda^2}{2} e^{-\lambda x_2} \mathrm{d}x_1 = \lambda^2 x_2 e^{-\lambda x_2} \qquad \text{for } x_2 \geq 0$$

i.e. $\Gamma(2, \lambda)$, as anticipated.

2.13 $\Pr(|X_1 - X_2| \leq x, \min(X_1, X_2) \leq y)$

$$= \Pr(|X_1 - X_2| \leq x, \min(X_1, X_2) \leq y, X_1 < X_2) + \ldots$$

$$= 2\Pr(X_1 - X_2 \leq x, X_2 \leq y, X_1 > X_2)$$

$$= 2\int_0^y f(u) \Pr(X_1 - X_2 \leq x, X_2 \leq y, X_1 > X_2 \mid X_2 = u)\,\mathrm{d}u$$

$$= 2\int_0^y f(u)\Pr(X_1 \leq x + u, X_1 > u)\,\mathrm{d}u = 2\int_0^y f(u)[F(x+u) - F(u)]\,\mathrm{d}u$$

Therefore the joint p.d.f. is

$$\frac{\mathrm{d}}{\mathrm{d}x}[2f(y)\{F(x+y) - F(y)\}] = 2f(y)\,f(x+y) = g(x, y), \text{ say.}$$

$$f(x) = \lambda e^{-\lambda x} \Rightarrow g(x, y) = 2\lambda e^{-\lambda y} \lambda e^{-\lambda(x+y)} = 2\lambda^2 e^{-2\lambda y} e^{-\lambda x}$$

$$\Rightarrow \text{independence}$$

Independence $\Rightarrow g(x, y) = 2f(y)f(x+y) = l(x)h(y)$ for some functions l and h.

Put $x = 0$, and let $K = l(0)$, then

$$2[f(y)]^2 = Kh(y) \Rightarrow h(y) = \frac{2}{K}[f(y)]^2$$

Also, $h(y) = \int_0^\infty g(x, y)\,dx = 2f(y)[1 - F(y)]$

Thus, $2f(y)[1-F(y)] = \frac{2}{K}[f(y)]^2$

$$1 - F(y) = f(y)/K$$

$$F(y) + \frac{dF}{dy}\Big/ K = 1$$

$F(y) = 1 + e^{-Ky}$, as required, Finally,

$$\Pr(X_1 + X_2 \leq 3\min(X_1, X_2) \leq 3b) = \Pr(U + 2V < 3V, V \leq b)$$

where $U = |X_1 - X_2|$ and $V = \min(X_1, X_2)$.

U, V are independent, from above, and $X_1 + X_2 = U + 2V$. Hence required probability $= \Pr(U < V \leq b)$

$$f_U(u) = \lambda e^{-\lambda u} \text{ (see Exercise 2.12) for some } \lambda > 0$$

(cf. Exercise 2.10), and $f_V(v) = 2\lambda e^{-2\lambda v}$, and using the independence property, we have required probability

$$\int_0^b \int_0^v \lambda e^{-\lambda u} 2\lambda e^{-2\lambda v}\,du\,dv = \frac{1}{3} - e^{-2\lambda b} + \frac{2}{3}e^{-3\lambda b}.$$

2.14 $f_X(x) = \frac{e^{-x/2}}{2^a}\frac{x^{a-1}}{\Gamma(a)}$ for $x \geq 0$

$f_Y(y) = \frac{e^{-y/2}y^{b-1}}{2^b\Gamma(b)}$ for $y \geq 0$

$$\begin{cases} S = X + Y & \frac{\partial s}{\partial x} = 1 \quad \frac{\partial s}{\partial y} = 1 \\ T = X/(X+Y) & \frac{\partial t}{\partial x} = \frac{y}{(x+y)^2} \quad \frac{\partial t}{\partial y} = -\frac{x}{(x+y)^2} \end{cases}$$

therefore $ST = X, Y = S(1 - T)$.

$$f_{ST}(s, t) = f_{XY}(x, y)(x+y) = \frac{e^{-((x+y)/2)}x^{a-1}y^{b-1}(x+y)}{2^{a+b}\Gamma(a)\Gamma(b)}$$

$$= \frac{se^{-s/2}(st)^{a-1}(s(1-t))^{b-1}}{2^{a+b}\Gamma(a)\Gamma(b)}$$

i.e. $$f_{ST}(s,t)=\frac{s^{a+b-1}e^{-s/2}t^{a-1}(1-t)^{b-1}\Gamma(a+b)}{2^{a+b}\Gamma(a+b)\Gamma(a)\Gamma(b)}.$$

We see that S and T are independent. S is $\chi^2_{2(a+b)}$ and T is $B_e(a, b)$.

2.15 (a) $$M_{N_1N_2}(\theta)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\exp[\theta x_1x_2]}{2\pi}\exp[-\tfrac{1}{2}(x_1^2+x_2^2)]\,dx_1\,dx_2$$

$$=\frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^{\infty}\exp[-x_1^2/2]$$

$$\times\left(\frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^{\infty}\exp[\theta x_1x_2-x_2^2/2]\,dx_2\right)dx_1$$

$$=\frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^{\infty}\exp[-x_1^2/2]\exp[+\tfrac{1}{2}\theta^2x_1^2]$$

$$\times\left(\frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^{\infty}\exp[-\tfrac{1}{2}(\theta x_1-x_2)^2]\,dx_2\right)dx_1$$

therefore $$=\frac{1}{\sqrt{(2\pi)}}\frac{1}{\sqrt{(1-\theta^2)}}\int_{-\infty}^{\infty}\exp[-y^2/2]\,dy$$

$$\left(\text{from setting } x_1=\frac{y}{\sqrt{(1-\theta^2)}}\right)$$

$$=\frac{1}{\sqrt{(1-\theta^2)}}$$

therefore $M_{N_1N_2+N_3N_4}(\theta)=\dfrac{1}{(1-\theta^2)}$

But this is the m.g.f. of a random variable with the Laplace distribution:

$$f_X(x)=\tfrac{1}{2}\exp[-|x|]\qquad\text{for } -\infty\le x\le\infty,$$

To see this:

$$M_X(\theta)=\frac{1}{2}\int_{-\infty}^{0}e^xe^{\theta x}dx+\frac{1}{2}\int_0^{\infty}e^{\theta x-x}dx$$

$$=\frac{1}{2(1+\theta)}-\frac{1}{2(\theta-1)},\qquad\text{for } \theta<1$$

$$=\frac{1}{(1-\theta^2)}$$

(This is the distribution of Y_1-Y_2 in Exercise 2.12.)

Hence

$$f_{|N_1N_2+N_3N_4|}(x) = e^{-x} \qquad \text{for } x \geq 0$$

(b) $C = N_1/N_2$. Also, set $D = N_1$, say, to give a 1–1 transformation from (N_1, N_2) to (C, D)

$$\frac{\partial c}{\partial n_1} = \frac{1}{n_2} \qquad \frac{\partial c}{\partial n_2} = -\frac{n_1}{n_2^2} \qquad \frac{\partial d}{\partial n_1} = 1 \qquad \frac{\partial d}{\partial n_2} = 0$$

Hence, $$f_{C,D}(c, d) = \frac{\exp[-\frac{1}{2}(n_1^2 + n_2^2)]}{2\pi} \left|\frac{n_2^2}{n_1}\right|$$

$$(d \equiv y, x \equiv c) = \frac{\exp[-\frac{1}{2}y^2(1+1/x^2)]}{2\pi} \left|\frac{y}{x^2}\right| = f_{X,Y}(x, y),$$

and now we form the marginal distribution of X:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,\mathrm{d}y = \int_{-\infty}^{0} f_{X,Y}(x, y)\,\mathrm{d}y + \int_{0}^{\infty} f_{X,Y}(x, y)\,\mathrm{d}y$$

$$= \int_{-\infty}^{0} \frac{1}{2\pi x^2}(-y)\exp[-(y^2/2)(1+1/x^2)]\,\mathrm{d}y$$

$$+ \int_{0}^{\infty} \frac{1}{2\pi x^2}\, y\exp[(-y^2/2)(1+1/x^2)]\,\mathrm{d}y$$

i.e. $$f_X(x) = \frac{1}{2\pi x^2}\frac{1}{(1+1/x^2)} + \frac{1}{2\pi x^2}\frac{1}{(1+1/x^2)} = \frac{1}{\pi(1+x^2)},$$

for $-\infty \leq x \leq \infty$,

i.e. X has the standard Cauchy distribution.

If X, Y are independent, identically distributed *exponential* random variables, with probability density function, e^{-x} for $x \geq 0$, then for $Z = X/Y$, $f_Z(z) = (1+z)^{-2}$ for $z \geq 0$, and $\mathscr{E}[Z] = \infty$.

2.16 $\phi(\mathbf{x}) = (2\pi)^{-p/2}|\mathbf{\Sigma}|^{-1/2}\exp\{-\frac{1}{2}\mathbf{x}'\mathbf{\Sigma}^{-1}\mathbf{x}\}\,\mathrm{d}\mathbf{x}$

Suppose, first of all, that $\mu = 0$:

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{z} \qquad \text{and the Jacobian is: } \|\mathbf{A}^{-1}\|$$

and so, $\phi(\mathbf{z}) = (2\pi)^{-p/2}|\mathbf{\Sigma}|^{-1/2}\exp\{-\frac{1}{2}(\mathbf{A}^{-1}\mathbf{z})'\mathbf{\Sigma}^{-1}\mathbf{A}^{-1}\mathbf{z}\}\,\|\mathbf{A}^{-1}\|$

i.e. $$\phi(\mathbf{z}) = (2\pi)^{-p/2}|\mathbf{\Sigma}|^{-1/2}\|\mathbf{A}\|^{-1}\exp\{-\frac{1}{2}(\mathbf{A}^{-1}\mathbf{z})'\mathbf{\Sigma}^{-1}\mathbf{A}^{-1}\mathbf{z}\}$$

$$= (2\pi)^{-p/2}|\mathbf{A}\mathbf{\Sigma}\mathbf{A}'|^{-1/2}\exp\{-\frac{1}{2}\mathbf{z}'(\mathbf{A}^{-1})'\mathbf{\Sigma}^{-1}\mathbf{A}^{-1}\mathbf{z}\}$$

$$= (2\pi)^{-p/2}|\mathbf{A}\mathbf{\Sigma}\mathbf{A}'|^{-1/2}\exp\{-\frac{1}{2}\mathbf{z}'(\mathbf{A}\mathbf{\Sigma}\mathbf{A}')^{-1}\mathbf{z}\}$$

i.e. $\mathbf{Z} = \mathbf{A}\mathbf{X}$ has an $N(\mathbf{0}, \mathbf{A}\mathbf{\Sigma}\mathbf{A}')$ distribution.