MATH764, Summer 2003. **Solutions**

1. [**(a) bookwork, (b) similar to homework.**]

(a) A Poisson random variable is a discrete random variable with probability mass function

$$p_k = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots,$$

where λ is a positive parameter which can be interpreted as the average number of events per unit time.

A Poisson random variable is a good approximation to a binomial random variable with n trials and probaibilty of success p if n is large and p is small and np moderate. The required relationship is

$$\lambda = np.$$

If $n \to \infty$, $p \to 0$ and $pn = \lambda$, the binomial probabilities tend to Poisson probabilities, i.e. for every fixed $k = 0, 1, 2, \ldots$

$$\binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!}\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{1}{k!}\lambda^k \left(1 - \frac{\lambda}{n}\right)^n \frac{n!}{n^k(n-k)!}\left(1 - \frac{\lambda}{n}\right)^{-k} \to \frac{\lambda^k}{k!}e^{-\lambda},$$

because (recall that k and λ are fixed)

$$\left(1 - \frac{\lambda}{n}\right)^n \to e^{-\lambda},$$

$$\frac{n!}{n^k(n-k)!} = \frac{n-k+1}{n}\frac{n-k+2}{n}\cdots\frac{n-1}{n}\frac{n}{n} \to 1,$$

$$\left(1 - \frac{\lambda}{n}\right)^{-k} \to 1.$$

(b) (i) There are

$$\lambda = \frac{8}{972} = 0.0082305$$

explosions per day on the average. In $t = 92$ days there are $t\lambda = 0.75720$ explosions on the average. Let X denote the number of explosions in a 92 day period. Then X is a Poisson random variable with rate $t\lambda$. It follows that

$$P(X = 0) = e^{-t\lambda} = e^{-0.75720} = 0.469.$$

Therefore the probability of at least one explosion is

$$P(X \geq 1) = 1 - P(X = 0) = 0.531,$$

i.e. around 53%.

(b) (ii) Let Y be the number of explosions in n days. We require that $P(Y \geq 1) = 0.95$. Equivalently,

$$P(Y=0) = 0.05; \qquad e^{-n\lambda} = 0.05; \qquad -n\lambda = \ln(0.05);$$

$$n = -\frac{\ln(0.05)}{0.0082305} = 363.98.$$

Thus the research should continue for 364 days in order to ensure that a supernova is observed with probability 0.95.

2. **[Not seen, but Bayes' rule was studied in depth.]**

(a) Let S and U denote the events "product will be successful" and "product will be unsuccessful". then

$$P(S) = 2/3; \qquad P(U) = 1/3.$$

Now, let X be the profit:

$$P(X = 1,500,000) = P(S) = 2/3; \quad P(X = -1,800,000) = P(U) = 1/3.$$

Thus $E[X] = 0.6667 \times 1,500,000 - 0.3333 \times 1,800,000 = 400,000$.

(b) Let S_p (U_p) denote the events "product is predicted to be successful (unsuccessful)". Then

$$P(S_p|S) = 0.8; \quad P(U_p|S) = 0.2; \quad P(S_p|U) = 0.3; \quad P(U_p|U) = 0.7.$$

The probabilities $P(S|S_p), P(U|S_p), P(S|U_p)$, and $P(U|U_p)$ required can be computed using the Bayes' rule:

$$P(S|S_p) = \frac{P(S_p|S)P(S)}{P(S_p|S)P(S) + P(S_p|U)P(U)} = \frac{0.8 \times 0.6667}{0.8 \times 0.6667 + 0.3 \times 0.3333} = 0.842;$$

$$P(U|S_p) = \frac{P(S_p|U)P(U)}{P(S_p|S)P(S) + P(S_p|U)P(U)} = \frac{0.3 \times 0.3333}{0.8 \times 0.6667 + 0.3 \times 0.3333} = 0.158;$$

$$P(S|U_p) = \frac{P(U_p|S)P(S)}{P(U_p|S)P(S) + P(U_p|U)P(U)} = \frac{0.2 \times 0.6667}{0.2 \times 0.6667 + 0.7 \times 0.3333} = 0.364;$$

$$P(U|U_p) = \frac{P(U_p|U)P(U)}{P(U_p|S)P(S) + P(U_p|U)P(U)} = \frac{0.7 \times 0.3333}{0.2 \times 0.6667 + 0.7 \times 0.3333} = 0.636.$$

(c) Let Y be the profit if the company follows the strategy described. Then

$$P(Y = 1,500,000) = P(S_p \text{ and } S) = P(S_p|S)P(S) = 0.533;$$

$$P(Y = -1,800,000) = P(S_p \text{ and } U) = P(S_p|U)P(U) = 0.1;$$

$$P(Y = 0) = P(U_p) = P(U_p|U)P(U) + P(U_p|S)P(S) = 0.3667.$$

Hence

$$E[Y] = 1,500,000 \times 0.533 - 1,800,000 \times 0.1 = 619,500.$$

Since $E[Y] - E[X] = 619,500 - 400,000 = 219,500 < 300,000$, the increase of the expected profit after carrying out the market survey is less than the price for that survey, and it is NOT worth carrying it out.

Obviously, the maximal acceptable price of such a survay is £219,500.

(d) Clearly, the above probability mass function for Y must be replaced with

$$P(Y = 1,500,000) = P(S_p|S)P(S) = 0.533;$$

$$P(Y = -1,800,000) = P(S_p|U)P(U) = (1 - \frac{k}{100}) \times 0.3333;$$

$$P(Y = 0) = P(U_p|U)P(U) + P(U_p|S)P(S) = 0.003333 \times k + 0.1334.$$

Hence, the increase of the expected profit is

$$E[Y] - E[X] = 1,500,000 \times 0.533 - 1,800,000 \times (0.3333 - 0.003333 \times k) - 400,000$$

$$= 6,000k - 200,000.$$

The survey is acceptable for £300,000 if $6,000k - 200,000 \geq 300,000$. Therefore, the minimal value of k equals

$$k = \frac{500,000}{6,000} = 83,33\%.$$

3. **[Similar to classwork and to homework.]**

(a)

(b) As usual, C is defined from the normalisation condition:

$$\frac{1}{C} = \int_0^1 \int_0^{1-x} x^2 y \, dy \, dx = \int_0^1 x^2 \frac{(1-x)^2}{2} dx = \frac{1}{2}\left[\frac{1}{3} - \frac{2}{4} + \frac{1}{5}\right] = \frac{1}{60}.$$

Thus $C = 60$.

(c)

$$f_X(x) = \int_0^{1-x} 60x^2 y dy = 60x^2 \int_0^{1-x} y dy = 60x^2 \left[\frac{(1-x)^2}{2}\right] = 30x^2(1-x)^2,$$

$$0 < x < 1.$$

$$f_Y(y) = \int_0^{1-y} 60x^2 y dx = 60y \int_0^{1-y} x^2 dx = 60y \left[\frac{(1-y)^3}{3}\right] = 20y(1-y)^3, \ \ 0 < y < 1.$$

(d)

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{60x^2 y}{30x^2(1-x)^3} = \frac{2y}{(1-x)^2}.$$

The above computation is valid for $0 < y < 1-x$. For y outside this range $f_{Y|X} \equiv 0$.

(e)

$$P(Y > 0.1|X = 0.5) = \int_{0.1}^{1-0.5} f_{Y|X}(y|0.5)dy = \int_{0.1}^{0.5} \frac{2y}{(1-0.5)^2} dy$$

$$= \frac{2}{0.25}\int_{0.1}^{0.5} y dy = \frac{2}{0.25}\frac{1}{2}\left[(0.5)^2 - (0.1)^2\right] = \frac{0.24}{0.25} = \frac{24}{25} = \mathbf{0.96}.$$

4. **[Similar to homework.]**

(a) Since

$$\frac{2}{\pi}\exp\left\{-\frac{x^2+y^2}{2}\right\} = \sqrt{\frac{2}{\pi}}\exp\left\{-\frac{x^2}{2}\right\} \times \sqrt{\frac{2}{\pi}}\exp\left\{-\frac{y^2}{2}\right\}$$

we can conclude that $f(x,y) = f_X(x)f_Y(y)$, meaning that X and Y are independent. The following straightforward calculations also receive the full mark. The marginal density of X is

$$f_X(x) = \int_0^\infty f(x,y)dy = \int_0^\infty \frac{2}{\pi}\exp\left\{-\frac{x^2+y^2}{2}\right\} dy$$

$$= \frac{2}{\pi}\exp\left\{-\frac{x^2}{2}\right\}\int_0^\infty \exp\left\{-\frac{y^2}{2}\right\} dy = \frac{2}{\pi}\exp\left\{-\frac{x^2}{2}\right\}\sqrt{2\pi}\int_0^\infty \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{y^2}{2}\right\} dy$$

$$= \frac{2}{\pi}\exp\left\{-\frac{x^2}{2}\right\}\sqrt{2\pi}\frac{1}{2} = \sqrt{\frac{2}{\pi}}\exp\left\{-\frac{x^2}{2}\right\},$$

because

$$\int_0^\infty \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{y^2}{2}\right\} dy = \frac{1}{2}\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{y^2}{2}\right\} dy = \frac{1}{2}.$$

Similarly,

$$f_Y(y) = \sqrt{\frac{2}{\pi}} \exp\left\{-\frac{y^2}{2}\right\}.$$

Since $f(x,y) = f_X(x)f_Y(y)$, we conclude that X and Y are independent.

(b) Solving the equations

$$u = x + 2y, \qquad v = x/y$$

for x and y, we obtain that the inverse transformation is

$$x = \frac{vu}{v+2}, \qquad y = \frac{u}{v+2}.$$

We compute the Jacobian J of the inverse transformation:

$$\frac{\partial x}{\partial u} = \frac{v}{v+2}; \quad \frac{\partial x}{\partial v} = \frac{2u}{(v+2)^2}; \quad \frac{\partial y}{\partial u} = \frac{1}{v+2}; \quad \frac{\partial y}{\partial v} = -\frac{u}{(v+2)^2};$$

$$J = -\left[\frac{vu}{(v+2)^3} + \frac{2u}{(v+2)^3}\right] = -\frac{(v+2)u}{(v+2)^3} = -\frac{u}{(v+2)^2}.$$

It follows that

$$\begin{aligned} f_{UV}(u,v) &= f\left(\frac{vu}{v+2}, \frac{u}{v+2}\right)\frac{u}{(v+2)^2} \\ &= \frac{2}{\pi}\exp\left\{-\frac{1}{2}\left[\frac{v^2u^2}{(v+2)^2} + \frac{u^2}{(v+2)^2}\right]\right\}\frac{u}{(v+2)^2} \\ &= \frac{2u}{\pi(v+2)^2}\exp\left\{\frac{-u^2(1+v^2)}{2(v+2)^2}\right\}, \qquad u, v > 0. \end{aligned}$$

(c) By the definition,

$$f_V(v) = \int_{-\infty}^{\infty} f_{UV}(u,v)du = \int_0^{\infty} \frac{2u}{\pi(v+2)^2}\exp\left\{\frac{-u^2(1+v^2)}{2(v+2)^2}\right\}du$$

$$= \left(\frac{u^2}{(v+2)^2} = t\right) = \frac{1}{\pi}\int_0^{\infty} e^{-t\frac{1+v^2}{2}}dt = \frac{2}{\pi(1+v^2)}, \quad v > 0.$$

(d) The conditional density

$$f_{U|V}(u|v) = \frac{f_{UV}(u,v)}{f_V(v)} = \frac{u(1+v^2)}{(v+2)^2}\exp\left\{\frac{-u^2(1+v^2)}{2(v+2)^2}\right\}$$

depends on v. Therefore U and V are dependent.

5. **[Bookwork, a similar problem with Poisson distribution was discussed in class.]**

MGF of RV X is defined as

$$M_X(t) = E[e^{tX}].$$

Properties:

(i)

$$E[X^r] = \left.\frac{d^r M_X(t)}{dt^r}\right|_{t=0}.$$

(ii) The MGF defines the distribution, i.e. if X and Y have the same MGF, then they have the same distribution.

(iii) If X has the MGF $M_X(t)$, then

$$M_{a+bX}(t) = e^{at} M_X(bt).$$

(iv) If X and Y are independent, then

$$M_{X+Y}(t) = M_X(t) M_Y(t).$$

(v) Suppose that

$$S = \sum_{i=1}^{N} X_i,$$

where X_i are iid RVs with the same MGF $M_X(t)$, and where N is independent of $\{X_i\}$ and has the MGF $M_N(t)$. Then

$$M_S(t) = M_N(\ln M_X(t)).$$

(a) For geometric RV, we have

$$M(t) = Ee^{tX} = \sum_{k=1}^{\infty} e^{tk} P(X = k) = \sum_{k=1}^{\infty} e^{tk} pq^{k-1} = pe^t \sum_{k=1}^{\infty} e^{t(k-1)} q^{k-1}$$

$$= pe^t \sum_{k=1}^{\infty} \left(e^t q\right)^{k-1} = pe^t \sum_{j=0}^{\infty} \left(e^t q\right)^j = \frac{pe^t}{1 - qe^t},$$

provided that $qe^t < 1$ i.e. $t < -\ln q$.

(b) Observe that

$$M'(t) = \frac{pe^t}{(1 - qe^t)^2}$$

and

$$M''(t) = \frac{pe^t(1 - qe^t)^2 + 2(1 - qe^t)qpe^t}{(1 - qe^t)^4}.$$

Consequently

$$EX = M'(0) = \frac{p}{(1 - q)^2} = \frac{1}{p}$$

and

$$EX^2 = M''(0) = \frac{p(1-q)^2 + 2(1-q)qp}{(1-q)^4} = \frac{2-p}{p^2}.$$

It follows that

$$\text{Var}X = EX^2 - (EX)^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

(c) According to the general properties of MGF

$$M_Z(t) = M_X(t)M_Y(t) = \left(\frac{pe^t}{1-qe^t}\right)^2, \qquad t < -\ln q.$$

Z has a negative binomial distribution with parameter $r = 2$ since Z =total number of trials up to (and including) the first 2 successes, in the series of independent Bernoulli(p) trials.

6. [**Similar to bookwork and homework.**]

(a) Observe that

$$\begin{aligned} P(|\bar{X} - \mu| < c) &= P(-c < \bar{X} - \mu < c) \\ &= P\left(-c < \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu) < c\right) \\ &= P\left(-c < \frac{\sigma}{\sqrt{n}} \frac{\sum_{i=1}^{n}(X_i - \mu)}{\sigma\sqrt{n}} < c\right) \\ &= P\left(-c < \frac{\sigma}{\sqrt{n}} Z_n < c\right), \end{aligned}$$

where Z_n is the normalized sum which has approximately standard normal distribution. We therefore have

$$\begin{aligned} P\left(-c < \frac{\sigma}{\sqrt{n}} Z_n < c\right) &= P\left(-\frac{c\sqrt{n}}{\sigma} < Z_n < \frac{c\sqrt{n}}{\sigma}\right) \\ &\approx \Phi\left(\frac{c\sqrt{n}}{\sigma}\right) - \Phi\left(-\frac{c\sqrt{n}}{\sigma}\right) = 2\Phi\left(\frac{c\sqrt{n}}{\sigma}\right) - 1. \end{aligned}$$

(b) Suppose now that $\sigma = 1$ and $c = 0.5$ and denote

$$\Phi = \Phi\left(\frac{c\sqrt{n}}{\sigma}\right).$$

The condition

$$P(|\bar{X} - \mu| < c) \geq 0.97$$

thus becomes $2\Phi - 1 \geq 0.97$, i.e. $\Phi \geq 0.985$. From the normal tables we find that $\Phi = 0.985$ means that

$$\frac{c\sqrt{n}}{\sigma} = 2.17.$$

Setting $\sigma = 1$ and $c = 0.5$, we get

$$\sqrt{n} = 2.17\frac{1}{0.5} = 4.34; \quad n = 18.84.$$

Thus we need at least 19 measurements.

(c) In the last case,

$$\frac{c\sqrt{n}}{\sigma} = 2.17,$$

where $c = 0.5$ and $n = 15$. Hence the maximal acceptable σ is

$$\sigma = \frac{0.5\sqrt{15}}{2.17} = 0.892.$$

7. **[Similar to examples discussed in class.]**

(a) According to the general theorem,

$$f_Z(z) = \int_{-\infty}^{\infty} |x| f_X(x) f_Y(xz) dx.$$

The densities of exponential random variables are

$$f_X(x) = \frac{1}{2}e^{-x/2}, \quad x \geq 0 \quad \text{and} \quad f_Y(y) = \frac{1}{2}e^{-y/2}, \quad y \geq 0.$$

Clearly, $f_Z(z) = 0$ if z is negative. If $z \geq 0$ then

$$f_Z(z) = \frac{1}{4}\int_0^{\infty} xe^{-x/2}e^{-xz/2}dx = (\text{by parts, } u = x/2, \ v = -\frac{e^{-x(z+1)/2}}{z+1})$$

$$= -\left.\frac{ue^{-u(z+1)}}{z+1}\right|_{u=0}^{u=\infty} + \int_0^{\infty} \frac{e^{-u(z+1)}}{z+1}du = -\left.\frac{e^{-u(z+1)}}{(z+1)^2}\right|_{u=0}^{u=\infty} = \frac{1}{(z+1)^2}.$$

(b) Now, since $\Gamma(1) = \Gamma(2) = 1$ we have: $f_Z(\cdot) = f_W(\cdot)$, where $m = n = 2$.

<u>Remark.</u> If someone remembers the χ^2 density

$$f_V(v) = \frac{(1/2)^{n/2}}{\Gamma(\frac{n}{2})}v^{\frac{n}{2}-1}e^{-v/2}. \quad v \geq 0$$

and notices that X and Y are just chi-square random variables with parameters $m = n = 2$ then he/she can argue like

"Z is the ratio of two independent chi-square random variables with $m = n = 2$;

hence $Z \sim F(2,2)$."
Such reasoning is worth full 15 marks.

(c) Reasoning similarly to (a) we obtain

$$f_Z(z) = \int_0^\infty xe^{-x}e^{-xz}dx = \int_0^\infty \frac{e^{-x(z+1)}}{z+1}dx = \frac{1}{(z+1)^2}, \quad z \geq 0,$$

i.e. the distribution of Z does not change.

(d) Since

$$\int_0^N zf_Z(z)dz = \int_0^N \frac{z\ dz}{(z+1)^2} = \int_0^N \frac{(z+1)dz}{(z+1)^2} - \int_0^N \frac{dz}{(z+1)^2} =$$

$$\ln(z+1)|_0^N + \left.\frac{1}{z+1}\right|_0^N = \ln(N+1) + \frac{1}{N+1} - 1 \to \infty \quad \text{as} \quad N \to \infty$$

we conclude that $E[Z]$ does not exist.