MATH764 January 2006 Exam Solutions

All questions similar to seen exercises except where marked as Bookwork (B) or Unseen (U).

1. (a) (i) $\Pr(X = 0) = \Pr(\text{First ball drawn is red}) = 1/2$
$\Pr(X = 1) = \Pr(\text{First white, second red}) = (1/2) \times (3/5) = 3/10$
$\Pr(X = 2) = \Pr(\text{First two white, third red}) = (1/2) \times (2/5) \times (3/4) = 3/20$
$\Pr(X = 3) = \Pr(\text{First three white}) = (1/2) \times (2/5) \times (1/4) = 1/20$

(ii) $E[X] = (1/2) \times 0 + (3/10) \times 1 + (3/20) \times 2 + (1/20) \times 3 = 3/4$
$\text{Var}[X] = E\left[X^2\right] - (E[X])^2 = ((1/2) \times 0^2 + (3/10) \times 1^2 + (3/20) \times 2^2 + (1/20) \times 3^2) - (3/4)^2 = (27/20) - (9/16) = 63/80 = 0.7875$

(b) (i) For drawing with replacement, the probability of drawing a red ball is 1/2 at each draw and successive draws are independent, so the probability of drawing y white balls followed by a red ball is $P(Y = y) = (1/2) \times (1 - (1/2))^y = (1/2)^{y+1}$ for $y = 0, 1, 2, \ldots$. That is, $Y + 1$ is geometrically distributed with success probability $p = 1/2$. Hence $E[Y] = (1/p) - 1 = 1$, $\text{Var}[Y] = (1 - p)/p^2 = 2$.

(ii) Now the event $V = v$ means that draw number $v + 1$ yields a red ball, while draws 1 to v yield 1 red, $v - 1$ white balls. The probability of this event is, for $v = 1, 2, 3, \ldots$,

$$P(V = v) = \left(\binom{v}{1}\left(\frac{1}{2}\right)^{v-1} \times \frac{1}{2}\right) \times \frac{1}{2} = v\left(\frac{1}{2}\right)^{v+1}.$$

$V+1$ has Negative Binomial distribution with parameters $\left(2, \frac{1}{2}\right)$, so that from (i) above, $E[V] = 2E[Y + 1] - 1 = 3$, $\text{Var}[V] = 2\text{Var}[Y] = 4$.

2. (a) For $t \geq 0$,

$$\begin{aligned}
F_T(t) &= \int_0^t \lambda e^{-\lambda u}\, du \\
&= \left[e^{-\lambda u}\right]_0^t \\
&= 1 - e^{-\lambda t} \qquad \text{for } t \geq 0 \\
F_T(t) &= 0 \qquad \text{for } t < 0
\end{aligned}$$

$$\begin{aligned}
E[T] &= \int_{-\infty}^{\infty} t f(t)\, dt \\
&= \int_0^{\infty} t\, \lambda e^{-\lambda t}\, dt \\
&= \left[-te^{-\lambda t}\right]_0^{\infty} + \int_0^{\infty} \lambda e^{-\lambda t}\, dt \\
&= 0 + \left[\frac{e^{-\lambda t}}{-\lambda}\right]_0^{\infty}
\end{aligned}$$

$$
\begin{aligned}
&= 0 + \frac{1}{\lambda} \\
&= \frac{1}{\lambda}
\end{aligned}
$$

Median m satsfies

$$
\begin{aligned}
F_T(m) &= 0.5 \\
1 - \mathrm{e}^{-\lambda m} &= 0.5 \\
-\lambda m &= \ln(0.5) = -\ln(2) \\
m &= \ln(2)/\lambda
\end{aligned}
$$

Lower quartile:

$$
\begin{aligned}
F_T(m) &= 0.25 \\
1 - \mathrm{e}^{-\lambda m} &= 0.25 \\
-\lambda m &= \ln(0.75) = -\ln(4/3) \\
m &= \ln(4/3)/\lambda
\end{aligned}
$$

Upper quartile:

$$
\begin{aligned}
F_T(m) &= 0.75 \\
1 - \mathrm{e}^{-\lambda m} &= 0.75 \\
-\lambda m &= \ln(0.25) = -\ln(4) \\
m &= \ln(4)/\lambda
\end{aligned}
$$

(b) (i) Have $0 \leq T < \infty$ and $Y = 1 - \mathrm{e}^{-2T}$, so that $0 \leq Y < 1$.

(ii) For $0 \leq y < 1$,

$$
\begin{aligned}
\Pr(Y \leq y) &= \Pr\left(1 - \mathrm{e}^{-2T} \leq y\right) \\
&= \Pr\left(\mathrm{e}^{-2T} \geq 1 - y\right) \\
&= \Pr\left(-2T \geq \ln\left(1 - y\right)\right) \\
&= \Pr\left(T \leq -(1/2)\ln(1-y)\right) \\
&= F_T\left(-(1/2)\ln(1-y)\right) \\
&= 1 - \mathrm{e}^{\lambda(1/2)\ln(1-y)} \\
&= 1 - \mathrm{e}^{2\ln(1-y)} \\
&= 1 - (1-y)^2 \\
F_Y(y) &= \begin{cases} 0 & y < 0 \\ 1 - (1-y)^2 & 0 \leq y < 1 \\ 1 & y \geq 1 \end{cases}
\end{aligned}
$$

(iii) For $0 \leq y < 1$,

$$
f_Y(y) = \frac{d}{dy}\left(1 - (1-y)^2\right)
$$

$$
\begin{aligned}
&= 2(1-y) \\
f_Y(y) &= \begin{cases} 2(1-y) & 0 \le y < 1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

(iv)

$$
\begin{aligned}
E[Y] &= \int_0^1 2y(1-y)\,dy = \left[y^2 - \frac{2y^3}{3}\right]_0^1 = 1 - \frac{2}{3} = \frac{1}{3} \\
E\left[Y^2\right] &= \int_0^1 2y^2(1-y)\,dy = \left[\frac{2y^3}{3} - \frac{y^4}{2}\right]_0^1 = \frac{2}{3} - \frac{1}{2} = \frac{1}{6} \\
\text{Var}[Y] &= \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{18}
\end{aligned}
$$

Median:

$$
\begin{aligned}
F_Y(m) &= 0.5 \\
1-(1-m)^2 &= 0.5 \\
(1-m)^2 &= 0.5 \\
1-m &= 1/\sqrt{2} \\
m &= 1-(1/\sqrt{2}) = 0.2929
\end{aligned}
$$

3. (a) $\text{Cov}[X,Y] = E\left[(X-E[X])(Y-E[Y])\right]$
$\text{Corr}[X,Y] = \text{Cov}[X,Y]/\sqrt{\text{Var}[X]\text{Var}[Y]}$
Correlation values lie between -1 and +1; positive correlation indicates that the two variables tend to increase together, negative correlation that as one increases, the other decreases; the larger the absolute value of correlation, the stronger the linear relationship. Correlation +1 and -1 indicate a perfect linear relationship between the two variables; correlation 0 indicates no linear relationship.

(b) (i) Marginal mass functions:
$\Pr(X=2) = 0.5$, $\Pr(X=4) = 0.5$
$\Pr(Y=0) = 0.3$, $\Pr(Y=1) = 0.3$, $\Pr(Y=2) = 0.4$
$E[X] = 0.5 \times 2 + 0.5 \times 4 = 3$; $E[Y] = 0.3 \times 0 + 0.3 \times 1 + 0.4 \times 2 = 1.1$
$\text{Var}[X] = \left(0.5 \times 2^2 + 0.5 \times 4^2\right) - 3^2 = 10 - 9 = 1$
$\text{Var}[Y] = \left(0.3 \times 0^2 + 0.3 \times 1^2 + 0.4 \times 2^2\right) - 1.1^2 = 1.9 - 1.21 = 0.69$

(ii) $\text{Cov}[X,Y] = (0.1 \times 0 \times 2 + 0.1 \times 1 \times 2 + 0.3 \times 2 \times 2 + 0.2 \times 0 \times 4 + 0.2 \times 1 \times 4$
$+0.1 \times 2 \times 4) - 3 \times 1.1 = 3 - 3.3 = -0.3$
$\text{Corr}[X,Y] = -0.3/\sqrt{1 \times 0.69} \approx -0.3612$

(iii) Correlation value indicates a weak/moderate negative relationship between X and Y, as can be seen from the joint mass function, where the smallest probability values of 0.1 are associated with (X,Y) pairs where either X and Y are both small or X and Y are both large.

(iv) $\text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y] + 2\,\text{Cov}[X,Y] = 1 + 0.69 - 2 \times 0.3 = 1.09$
$\text{Var}[X-Y] = \text{Var}[X] + \text{Var}[Y] - 2\,\text{Cov}[X,Y] = 1 + 0.69 + 2 \times 0.3 = 2.29$

4. (a)

$$\begin{aligned}
\int f(x,y)dydx &= k\int_{x=0}^{1}\int_{y=0}^{1} x(1-x)+y(1-y)dydx \\
&= k\left(\left[\frac{x^2}{2}-\frac{x^3}{3}\right]_{x=0}^{1}+\left[\frac{y^2}{2}-\frac{y^3}{3}\right]_{y=0}^{1}\right) \quad = \quad k\left(\frac{1}{6}+\frac{1}{6}\right) \quad = \quad \frac{k}{3}
\end{aligned}$$

So $k=3$.

Marginals:

$$\begin{aligned}
f_X(x) &= \int_{y=0}^{1} f(x,y)dy \quad = \quad \int_{y=0}^{1} 3\left(x(1-x)+y(1-y)\right)dy \\
&= 3\left[x(1-x)y+\frac{y^2}{2}-\frac{y^3}{3}\right]_{y=0}^{1} \\
&= 3\left(x(1-x)+\frac{1}{6}\right) \quad = \quad 3x(1-x)+\frac{1}{2} \\
\text{By symmetry, } f_Y(y) &= 3y(1-y)+\frac{1}{2}
\end{aligned}$$

(b) Inverting the transformation,

$$\begin{aligned}
x=(u+v)/2, &\qquad \frac{dx}{du}=\frac{1}{2}, \qquad \frac{dx}{dv}=\frac{1}{2}, \\
y=(v-u)/2, &\qquad \frac{dy}{du}=-\frac{1}{2}, \qquad \frac{dy}{dv}=\frac{1}{2}, \\
&\qquad \frac{\partial(x,y)}{\partial(u,v)}=\left|\begin{array}{cc} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2}\end{array}\right|=\frac{1}{2}
\end{aligned}$$

$$\begin{aligned}
f_{U,V}(u,v) &= f_{X,Y}((u+v)/2,(v-u)/2)\left|\frac{\partial(x,y)}{\partial(u,v)}\right| \\
&= 3\left(\frac{1}{2}(u+v)\left(1-\frac{1}{2}(u+v)\right)+\frac{1}{2}(v-u)\left(1-\frac{1}{2}(v-u)\right)\right)\times\frac{1}{2} \\
&= (3/8)\left((u+v)(2-(u+v))+(v-u)(2-(v-u))\right) \\
&= (3/8)\left(2(u+v+v-u)-(u+v)^2-(v-u)^2\right) \\
&= (3/8)\left(4v-2u^2-2v^2\right) \\
&= (3/2)v-(3/4)\left(u^2+v^2\right)
\end{aligned}$$

Region of positive density:

5. (a) Differentiating $G_U(s) = E\left[s^U\right]$ with respect to s,

$$G_U'(s) = E\left[Us^{U-1}\right] \qquad G_U''(s) = E\left[U(U-1)s^{U-2}\right]$$
$$\text{so that } G_U'(1) = E[U] \qquad G_U''(1) = E[U(U-1)]$$

and so

$$G''(1) + G'(1) - (G_U'(1))^2 = E[U(U-1)] + E[U] - (E[U])^2 = E\left[U^2\right] - (E[U])^2 = \text{Var}[U]$$

as required.

(b)

$$\begin{aligned} G_V(s) &= \sum_{k=1}^{\infty} s^k \left(\frac{1}{2^k k \ln(2)}\right) &&= \frac{1}{\ln(2)} \sum_{k=1}^{\infty} \frac{1}{k}\left(\frac{s}{2}\right)^k \\ &= \frac{1}{\ln(2)} \ln\left(\frac{1}{1-(s/2)}\right) &&= \ln\left(\frac{2}{2-s}\right) \Big/ \ln(2) \\ &= \frac{\ln(2) - \ln(2-s)}{\ln(2)} &&= 1 - \frac{\ln(2-s)}{\ln(2)} \end{aligned}$$

(c) Differentiating,

$$\begin{aligned} G_V'(s) &= -\left(\frac{1}{2-s}\right) \times (-1) \times \frac{1}{\ln(2)} = \frac{1}{\ln(2)}\left(\frac{1}{2-s}\right) \Rightarrow G'(1) = \frac{1}{\ln(2)} \\ G_V''(s) &= \frac{1}{\ln(2)}\left(\frac{1}{2-s}\right)^2 \Rightarrow G''(1) = \frac{1}{\ln 2} \end{aligned}$$

so that

$$E[V] = \frac{1}{\ln(2)}, \qquad \text{Var}[V] = \frac{1}{\ln(2)} + \frac{1}{\ln(2)} - \left(\frac{1}{\ln(2)}\right)^2 = \frac{2\ln(2) - 1}{(\ln(2))^2}$$

6. (a) Since the X_i are independent, identically distributed, and have finite variance, we can apply the CLT, and hence approximately have $\bar{X} \sim N(9, 9/n)$.

(b) Using the CLT approximation, we require to find n such that $P\left(\bar{X} > 9.05\right) = 0.01$, which is approximately equivalent to

$$\begin{aligned} P\left(Z > \frac{9.05-9}{\sqrt{9/n}}\right) &= 0.01 \\ \Rightarrow P\left(Z < \frac{9.05-9}{\sqrt{9/n}}\right) &= 0.99 \\ \Rightarrow \frac{0.05}{\sqrt{9/n}} &= 2.33 \\ 0.05 &= 2.33\sqrt{9/n} = 6.99/\sqrt{n} \\ \Rightarrow \sqrt{n} &= 6.99/0.05 = 139.8 \\ \Rightarrow n &= 19544.04 \end{aligned}$$

So need $n = 19545$.

(c) By CLT,

$$\begin{aligned}
P\left(|\bar{X}-9|<c\right) &= P\left(\frac{|\bar{X}-9|}{\sqrt{9/n}}<\frac{c}{\sqrt{9/n}}\right) \approx P\left(|Z|<\frac{c}{\sqrt{9/n}}\right) \\
&= P\left(-\frac{c\sqrt{n}}{3}<Z<\frac{c\sqrt{n}}{3}\right) \\
&= P\left(Z<\frac{c\sqrt{n}}{3}\right)-\left(1-P\left(Z<\frac{c\sqrt{n}}{3}\right)\right) \\
&= 2\Phi\left(\frac{c\sqrt{n}}{3}\right)-1.
\end{aligned}$$

(d) Next, have $\sum_{i=1}^{5} X_i \sim \text{Poisson}(45)$, and so

$$\begin{aligned}
P\left(|\bar{X}-9|<0.3\right) &= P\left(\left|\sum_{i=1}^{5} X_i - 45\right| < 1.5\right) \\
&= P\left(43.5<\sum_{i=1}^{5} X_i<46.5\right) \\
&= \mathrm{e}^{-45}\left(\frac{45^{44}}{44!}+\frac{45^{45}}{45!}+\frac{45^{46}}{46!}\right) \\
&= \mathrm{e}^{-45}\frac{45^{44}}{44!}\left(1+1+\frac{45}{46}\right) \\
&= 0.1768
\end{aligned}$$

(e) Using approximation of part (c),

$$\begin{aligned}
P\left(|\bar{X}-9|<0.3\right) &\approx 2\Phi\left(\frac{0.3\sqrt{5}}{3}\right)-1 \\
&= 2\Phi(0.22)-1 = 2\times 0.5871-1 = 0.1742
\end{aligned}$$

(f) Comment: the approximation is pretty good, in that the exact and approximate answers differ by only 0.0026, in spite of the fact that n is small. This is because in approximating Poisson by normal, we need the mean of the Poisson to be reasonably large. Specifically, the approximation is certainly good for mean ≥ 10. In this case the mean is 45, which is certainly large enough.

7. (a) For $y \geq 0$,

$$\begin{aligned}
\Pr(Y\leq y) &= \Pr\left(Z^2\leq y\right) \\
&= \Pr\left(-\sqrt{y}\leq Z\leq\sqrt{y}\right) \\
&= 2\Pr\left(0\leq Z\leq\sqrt{y}\right) \\
&= 2\int_0^{\sqrt{y}}\frac{1}{\sqrt{2\pi}}\exp\left(-z^2/2\right)dz \\
\text{so } f_Y(y) &= \frac{d}{dy}\Pr\left(Y\leq y\right) \\
&= \frac{2}{\sqrt{2\pi}}\exp\left(-y/2\right)\times(1/2)y^{-1/2} \\
&= \frac{1}{\sqrt{2\pi y}}\exp(-y/2) \qquad (y\geq 0)
\end{aligned}$$

(b) $M_Y(t) = \left(\dfrac{1}{1-2t}\right)^{1/2}$.

(c) If $Z_1, Z_2, \ldots, Z_n$ are independent standard normal random variables, then the distribution of $V = Z_1^2 + \cdots + Z_n^2$ is called the chi-squared distribution with n degrees of freedom.

Hence

$$M_V(t) \quad = \quad E\left[\mathrm{e}^{-tV}\right] \quad = \quad E\left[\mathrm{e}^{-t(Y_1+\cdots+Y_n)}\right] \quad = \quad \left(E\left[\mathrm{e}^{-tY}\right]\right]^n$$

where $Y_1, \ldots, Y_n$ are independent χ_1^2 random variables, so each have the moment generating function $M_Y(t)$ of part (b), and hence $M_V(t) = \left(\dfrac{1}{1-2t}\right)^{n/2}$.

From the given expression for mgf of Gamma distribution, have $\lambda_n = \frac{1}{2}$ and $\alpha_n = \frac{n}{2}$.

(d) First, mgf of W_n is

$$M_n(t) \quad = \quad E\left[\mathrm{e}^{tW_n}\right] \quad = \quad E\left[\mathrm{e}^{tV_n/n}\right] \quad = \quad M_{V_n}\left(\frac{t}{n}\right) \quad = \quad \left(\frac{1}{1-\frac{2t}{n}}\right)^{n/2}.$$

Taking the limit,

$$\begin{aligned} M(t) \quad &= \quad \lim_{n\to\infty} M_n(t) \quad = \quad \lim_{n\to\infty}\left(1-\frac{2t}{n}\right)^{-n/2} \quad = \quad \lim_{n\to\infty}\left(\left(1-\frac{2t}{n}\right)^n\right)^{-/12} \\ &= \quad \left(\mathrm{e}^{-2t}\right)^{-1/2} \quad = \quad \mathrm{e}^t \end{aligned}$$

Hence in the limit as $n \to \infty$, the random variable W_n converges to a random variable W with $E\left[\mathrm{e}^{tW}\right] = \mathrm{e}^t$, so that $W = 1$ with probability 1.