MATH264 January 2006 Exam Solutions

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

1. (a) If $X \sim \text{Bin}(n,p)$, then for $x = 0, 1, \ldots, n$,

$$\begin{aligned}
\Pr(X = x) &= \frac{n!}{x!\,(n-x)!} p^x (1-p)^{n-x} \\
&= n(n-1)\ldots(n-x+1)\,\frac{p^x}{x!}\,(1-p)^{n-x} \\
&= \left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\ldots\left(\frac{n-x+1}{n}\right)\frac{(np)^x}{x!}\,(1-p)^{n-x} \\
&= 1\left(1-\frac{1}{n}\right)\ldots\left(1-\frac{x-1}{n}\right)\frac{(np)^x}{x!}\,(1-p)^{n-x}
\end{aligned}$$

Now suppose that $n \to \infty$ and $p \to 0$ in such a way that $np \to \lambda$. Then for fixed x,

$$\lim_{\substack{n\to\infty \\ p\to 0}} \Pr(X = x) = \lim_{\substack{n\to\infty \\ p\to 0}} 1\left(1-\frac{1}{n}\right)\ldots\left(1-\frac{x-1}{n}\right)\frac{\lambda^x}{x!}\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^x}$$

Now as $n \to \infty$,

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

For each $s = 1, 2, \ldots, x-1$, and for $s = \lambda$,

$$\left(1-\frac{s}{n}\right) \to 1$$

Hence

$$\begin{aligned}
\Pr(X = x) &\to 1 \times 1 \times \ldots \times 1\,\frac{\lambda^x}{x!}\,\frac{\mathrm{e}^{-\lambda}}{1} \\
&= \frac{\lambda^x}{x!}\,\mathrm{e}^{-\lambda}
\end{aligned}$$

which is the Poisson(λ) probability mass function.

(b) Poisson provides good approximation to binomial for $n > 50$ with $np < 5$.

(c) Denoting by X the number with disease in the sample, then $X \sim \text{Bin}(400, 0.0002)$, since disease status of distinct individuals independent and in random sample probability of having disease is the same for each individual, being $1/5000 = 0.0002$. Approximating X by Poisson(0.08),

$$\begin{aligned}
\Pr(X \le 4) &= \mathrm{e}^{-0.08}\left(1 + \frac{0.08}{1!} + \frac{0.08^2}{2!} + \frac{0.08^3}{3!} + \frac{0.08^4}{4!}\right) \\
&= \mathrm{e}^{-0.08}\,(1 + 0.08 + 0.0032 + 0.000085333 + 0.000001706) \\
&= 1.08328704 \times \mathrm{e}^{-0.08} \\
&= 0.999999974
\end{aligned}$$

Poisson approximation is appropriate because number of trials $n = 400$ is large, success probability $p = 0.0002$ is small. More specifically, $n > 50$ and $np = 0.08 < 5$.

2. (a) $\Pr(X = 0) = \Pr(\text{First ball drawn is red}) = 1/2$
$\Pr(X = 1) = \Pr(\text{First white, second red}) = 1/2 \times (2/3) = 1/3$
$\Pr(X = 2) = \Pr(\text{First white, second white}) = 1/2 \times (1/3) = 1/6$

(b) $E[X] = (1/2) \times 0 + (1/3) \times 1 + (1/6) \times 2 = 2/3$
$\text{Var}[X] = E\left[X^2\right] - (E[X])^2 = \left((1/2) \times 0^2 + (1/3) \times 1^2 + (1/6) \times 2^2\right) - (2/3)^2 = 1 - (4/9) = 5/9$

(c) Letting $S = \sum_{i=1}^{100} X_i$, then S is approximately Normal, mean 200/3, variance 500/9.

$$\begin{aligned}
\Pr(64 \le S \le 81) &= \Pr(63.5 \le S \le 81.5) \\
&\approx \Pr\left(\frac{63.5 - (200/3)}{\sqrt{500/9}} \le S \le \frac{81.5 - (200/3)}{\sqrt{500/9}}\right) \\
&= \Pr(-0.42 \le Z \le 1.99) \\
&= \Pr(Z \le 1.99) - (1 - \Pr(Z \le 0.42)) \\
&= 0.9767 - (1 - 0.6628) \\
&= 0.6395
\end{aligned}$$

(d) 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 $y = 0, 1, 2, \ldots$.

3. (a) For $t \ge 0$,

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

$$\begin{aligned}
E[T] &= \int_{-\infty}^{\infty} t f(t)\, dt \\
&= \int_0^{\infty} t\, \lambda \mathrm{e}^{-\lambda t}\, dt \\
&= \left[-t\mathrm{e}^{-\lambda t}\right]_0^{\infty} + \int_0^{\infty} \lambda \mathrm{e}^{-\lambda t}\, dt \\
&= 0 + \left[\frac{\mathrm{e}^{-\lambda t}}{-\lambda}\right]_0^{\infty} \\
&= 0 + \frac{1}{\lambda} \\
&= \frac{1}{\lambda}
\end{aligned}$$

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

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

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

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

$$\begin{aligned}
f_Y(y) &= \frac{d}{dy}\left(1-(1-y)^2\right) \\
&= 2(1-y) \\
f_Y(y) &= \begin{cases} 2(1-y) & 0 \le y < 1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

(iv)

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

4. (a) $\mathrm{Cov}[X,Y] = E\left[(X - E[X])(Y - E[Y])\right]$
$\mathrm{Corr}[X,Y] = \mathrm{Cov}[X,Y]/\sqrt{\mathrm{Var}[X]\mathrm{Var}[Y]}$

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

(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.

5. (a)

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

So $k=3$.

Marginals:

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

(b) (i) Inverse transformation is $x=(u+v)/2$, $y=(v-u)/2$, so

$$\begin{aligned}\frac{dx}{du} &= \frac{1}{2}, & \frac{dx}{dv} &= \frac{1}{2}, \\ \frac{dy}{du} &= -\frac{1}{2}, & \frac{dy}{dv} &= \frac{1}{2}, \\ \frac{\partial(x,y)}{\partial(u,v)} &= \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{vmatrix} = \frac{1}{2}\end{aligned}$$

(ii)

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

(iii) Region of positive density:

6. (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

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

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

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

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

$$\begin{aligned} \Pr(Y \leq y) &= \Pr\left(Z^2 \leq y\right) \\ &= \Pr\left(-\sqrt{y} \leq Z \leq \sqrt{y}\right) \\ &= 2\Pr\left(0 \leq Z \leq \sqrt{y}\right) \end{aligned}$$

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

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

(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

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

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

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