Stat 331/Elec 331, Solutions to Final Exam

1. Let the cdf’s and pdf’s of Y and Z be denoted by \( F_Y, F_Z, f_Y \) and \( f_Z \) respectively. Note that the range of \( Y \) is \((1, \infty)\) and the range of \( Z \) is \((0, \infty)\). We get

\[
F_Y(x) = P(Y \leq x) = P(e^X \leq x) = P(X \leq \log x)
\]

\[
= F_X(\log x) = 1 - e^{-\log x} = 1 - \frac{1}{x}, \quad x > 1
\]

and hence

\[
f_Y(x) = \frac{d}{dx} F_Y(x) = \frac{1}{x^2}, \quad x > 1.
\]

Further,

\[
F_Z(x) = P(Z \leq x) = P(\frac{1}{X} \leq x) = P(X \geq \frac{1}{x})
\]

\[
= 1 - F_X(\frac{1}{x}) = 1 - e^{-1/x}
\]

and hence

\[
f_Z(x) = \frac{d}{dx} F_Z(x) = \frac{1}{x^2} e^{-1/x}, \quad x > 0.
\]

2a. Since \( C = 2X + 2Y \)

\[
E[C] = 2E[X] + 2E[Y] = 30
\]

\[
E[C | X = x] = 2x + 2E[Y|X = x]
\]

where

\[
E[Y|X = x] = 0.8 \frac{\sqrt{0.04}}{\sqrt{0.01}} (x - 5) = 1.6x + 2
\]

and hence

\[
E[C | X = x] = 5.2x + 4.
\]
b. Since $A = XY$ we get

$$E[A] = E[XY] = Cov[X, Y] + E[X]E[Y] = 0.8\sqrt{0.01 \cdot 0.04} + 5 \cdot 10 = 50.016$$

$$E[A|X = x] = xE[Y|X = x] = x(1.6x + 2) = 1.6x^2 + 2x.$$ 

c. The probability that a plate is useful is

$$P(29 \leq C \leq 31) = \Phi \left( \frac{31 - 30}{\sqrt{0.328}} \right) - \Phi \left( \frac{29 - 30}{\sqrt{0.328}} \right) = \Phi(1.75) - (1 - \Phi(1.75))$$

$$= 2\Phi(1.75) - 1 = 0.92$$

d. Let $X$ = the number of useful plates. Then $X$ has a binomial distribution with $n = 10$ and $p = P(\text{useful}) = 0.92$. Hence

$$P(X \geq 9) = P(X = 9) + P(X = 10) = \binom{10}{9} 0.92^9 0.08 + 0.92^{10} = 0.81.$$ 

3a. The likelihood function is

$$L(a) = \prod_{k=1}^{n} f(X_k) = \prod_{k=1}^{n} aX_k e^{-\frac{aX_k^2}{2}}$$

$$= a^n \prod_{k=1}^{n} X_k e^{-\frac{aX_k^2}{2}}.$$ 

Taking the logarithm gives

$$\log L(a) = n \log a - \frac{a}{2} \sum_{k=1}^{n} X_k^2 + \log(\prod_{k=1}^{n} X_k)$$

and differentiation with respect to $a$ gives

$$\frac{d}{da} \log L(a) = \frac{n}{a} - \frac{1}{2} \sum_{k=1}^{n} X_k^2.$$ 

Setting this equal to 0 gives the MLE
\[ \hat{a} = \frac{2n}{\sum_{k=1}^{n} X_k^2} . \]

b. Let \( F \) denote the cdf of \( X \). Then \( F^{-1}(U) \) has the same distribution as \( X \). First find \( F \):

\[
F(x) = \int_{0}^{x} f(t) dt = \left[ -e^{-\frac{u^2}{2}} \right]_0^x = 1 - e^{-\frac{x^2}{2}}
\]

which gives

\[ F^{-1}(x) = \sqrt{-\log(1 - x)} \]

and hence \( \sqrt{-\log(1 - U)} \) has the desired distribution.

4. The rates are

From 0 to 1: \( \alpha(1 - p) \)
From 0 to 2: \( \alpha p \)
From 1 to 2: \( \alpha(1 - p) \)
From 1 to 3: \( \alpha p \)
From 2 to 3: \( \alpha(1 - pq) \)

From 3 to 2: \( \beta \)
From 2 to 1: \( \beta \)
From 1 to 0: \( \beta \)

Note that, if the system is in state 2, it can go to state 3 in two ways: a single individual arrives, which has rate \( \alpha(1 - p) \) and a pair arrives and one of them joins, which has rate \( \alpha p(1 - q) \). Hence the total rate is \( \alpha(1 - p) + \alpha p(1 - q) = \alpha(1 - pq) \).

b. First note that if the system is in state 3, all arriving customers are lost. If the system is in state 2, half of all customers arriving in pairs are immediately lost, and of the other half, the proportion \( q \) is lost due to deciding not to stay. Thus, in state 2, the proportion lost is \( p_2^1 + p_2^2 q = p_2(1 + q) \) Hence, the expected proportion of lost customers is
\[ \pi_3 + \pi_2 p \frac{1}{2}(1 + q). \]

c. With \( \alpha = \beta \) and \( p = q = 1/2 \) the balance equations become

State 0: \( \alpha \pi_0 = \alpha \pi_1 \Rightarrow \pi_1 = \pi_0 \)

State 1: \( (\alpha + \alpha \frac{1}{2} + \alpha \frac{1}{2})\pi_1 = \alpha \frac{1}{2} \pi_0 + \alpha \pi_2 \Rightarrow \pi_2 = 2\pi_1 - \frac{1}{2} \pi_0 = \frac{3}{2} \pi_0. \)

State 3: \( \alpha \pi_3 = \alpha \frac{1}{2} \pi_1 + \alpha \frac{3}{2} \pi_2 \Rightarrow \pi_3 = \frac{13}{8} \pi_0. \)

The condition \( \sum \pi_k = 1 \) gives

\[ \pi_0(1 + 1 + \frac{3}{2} + \frac{13}{8}) = 1 \Rightarrow \pi_0 = \frac{8}{41} \]

and hence the stationary distribution is \( (8/41, 8/41, 12/41, 13/41) \).

5. The plots (b), (c), (d) and (e) are from bivariate normal distributions and the correlation coefficients are 0, -0.99, 0.5 and 0 respectively. The two plots that have \( \rho = 0 \) differ in that that in (e) the variances are the same and in (b) the variance of \( Y \) is higher than that of \( X \).

Plot (a) is uniform on a circle, plot (f) uniform on an ellips. In plot (g), \( X \) is normal but \( Y \) is discrete (note the horizontal lines) and in (h), \( X \) is normal and \( Y \) is normal with mean \( X^3 \) (you can imagine the curve \( y = x^3 \) centered in the plot).

6. First note that

\[ W = \begin{cases} \frac{L}{L - 30} & \text{if } L \leq 30 \\ L - 30 & \text{if } L \geq 30 \end{cases} \]

This means that \( W \) is a function of \( L \), \( W = g(L) \) say, and we can use the formula \( E[g(L)] = \int g(x)f_L(x)dx \) where \( f_L \) is the pdf of \( L \). Let \( F_L \) be the cdf of \( L \) to obtain
\[
E[W] = \int_{-\infty}^{30} x f_L(x)dx + \int_{30}^{\infty} (x - 30)f_L(x)dx
\]
\[
= \int_{-\infty}^{30} x f_L(x)dx + \int_{30}^{\infty} (x f_L(x)dx - 30 \int_{30}^{\infty} f_L(x)dx
\]
\[
= \int_{-\infty}^{\infty} x f_L(x)dx - 30(1 - F_L(30))dx = E[L] + 30F_L(30) - 30
\]
\[
= \mu + 30\Phi(30 - \mu) - 30
\]

since \( L \sim N(\mu, 1) \). This is a function of \( \mu \) that we want to minimize, so differentiation with respect to \( \mu \) gives

\[
\frac{d}{d\mu}(\mu + 30\Phi(30 - \mu) - 30) = 1 - 30\varphi(30 - \mu) = 1 - \frac{30}{\sqrt{2\pi}} e^{-\frac{1}{2}(30-\mu)^2}.
\]

Setting this equal to 0 gives

\[
\mu = 30 + \sqrt{2 \log\left(\frac{30}{\sqrt{2\pi}}\right)} = 32.2.
\]

7. The elements yttrium, erbium, terbium and ytterbium are all named for the village Ytterby outside Stockholm, Sweden.