Stat 330 Assignment 3 Solutions

Mandy Yao

October 2020

3.1.29

X_1 and X_2 are independent, with $X_1 \sim Bin(n_1, \frac{1}{2})$ and $X_2 \sim Bin(n_2, \frac{1}{2})$. We want to show that $Y = X_1 - X_2 + n_2 \sim Bin(n_1 + n_2, \frac{1}{2})$. We have that

$$M_Y(t) = E(e^{tY}) = E(e^{t(X_1 - X_2 + n_2)}) = E(e^{tX_1})E(e^{-tX_2})E(e^{tn_2}) = E(e^{tX_1})E(e^{-tX_2})e^{tn_2}$$

$$= e^{tn_2}\left(\frac{1}{2} + \frac{1}{2}e^t\right)^{n_1}\left(\frac{1}{2} + \frac{1}{2}e^{-t}\right)^{n_2} = \left(\frac{1}{2} + \frac{1}{2}e^t\right)^{n_1}\left(\frac{1}{2}e^t + \frac{1}{2}\right)^{n_2} = \left(\frac{1}{2} + \frac{1}{2}e^t\right)^{n_1+n_2}.$$

By the uniqueness of MGFs, $Y \sim Bin(n_1 + n_2, \frac{1}{2})$.

3.5.6

We have that

$$E(e^{tUV}) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{tuv} f_{U,V}(u, v)dudv.$$

Since U and V are independent,

$$f_{U,V}(u, v) = f_U(u)f_V(v) = \frac{1}{2\pi}e^{-(u^2+v^2)/2}.$$

This means that we need to find

$$E(e^{tUV}) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{1}{2\pi}e^{tuv-u^2/2-v^2/2}dudv.$$

We can transform the exponent:

$$tuv - \frac{u^2}{2} - \frac{v^2}{2} = -\frac{1}{2}(u^2 - 2tuv + v^2) = -\frac{1}{2(1-t^2)}(u^2(1-t^2) - 2tuv(1-t^2) + v^2(1-t^2))$$

$$= -\frac{1}{2(1-t^2)}\Big[\left(\frac{u}{1/\sqrt{1-t^2}}\right)^2 - 2t\left(\frac{u}{1/\sqrt{1-t^2}}\right)\left(\frac{v}{1/\sqrt{1-t^2}}\right) + \left(\frac{v}{1/\sqrt{1-t^2}}\right)^2\Big].$$

Denote

$$-\frac{1}{2(1-t^2)}\Big[\left(\frac{u}{\sigma_1}\right)^2 - 2t\left(\frac{u}{\sigma_1}\right)\left(\frac{v}{\sigma_2}\right) + \left(\frac{v}{\sigma_2}\right)^2\Big] = -\frac{q}{2},$$

where $\sigma_1 = \sigma_2 = \frac{1}{\sqrt{1-t^2}}$. With parameters $\mu_1 = \mu_2 = 0, \sigma_1 = \sigma_2 = (1-t^2)^{-1/2}, \rho = t$, the integral of the bivariate normal pdf is

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-t^2}}e^{-q/2}dudv = 1$$

Then we have that

$$E(e^{tUV}) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{1}{2\pi} e^{-q/2} du dv = \sigma_1\sigma_2\sqrt{1-t^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-t^2}} e^{-q/2} du dv$$

$$= \sigma_1\sigma_2\sqrt{1-t^2} = \frac{1}{\sqrt{1-t^2}}\frac{1}{\sqrt{1-t^2}}\sqrt{1-t^2} = (1-t^2)^{-1/2}.$$

Since $1-t^2 > 0$, we require that $-1 < t < 1$.

3.5.6 (7th Edition of Textbook)

We have that $(X, Y) \sim BN(\underline{\mu}, \Sigma)$, with $\underline{\mu} = \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix} = \begin{pmatrix}20\\ 40\end{pmatrix}$, $\Sigma = \begin{pmatrix}9 & 3.6\\ 3.6 & 4\end{pmatrix}$. From this, we get that

$$E(Y|X=22) = \mu_1 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1) = 40 + \left(\frac{6}{10}\right)\left(\frac{2}{3}\right)(22-20) = 40.8$$

Then the shortest interval for which $P(Y|X=22) = 0.90$ is

$$P(Y|X=22) = E(Y|X=22) \pm 1.645(2)\sqrt{1-0.6^2} = 40.8 \pm 2.6 = (38.2, 43.4)$$

3.6.11

We have that $T = W/\sqrt{V/r}$, W and V are independent, $W \sim N(0,1)$, and $V \sim \chi^2_{(r)}$. We want to show that $T^2 \sim F(1,r)$. First, we can see that

$$T^2 = (W/\sqrt{V/r})^2 = \frac{W^2}{V/r}.$$

By theorem 3.4.1 in the textbook, $W^2 \sim \chi^2_{(1)}$. Therefore, since W^2 and V are independent,

$$T^2 = \frac{W^2/1}{V/r} \sim F(1,r).$$

3.7.13

We have that $g(\theta) \sim \Gamma(\alpha, \beta)$, so

$$f(x,\theta) = f(x|\theta)g(\theta) = \theta\tau x^{\tau-1}e^{-\theta x^\tau}\frac{1}{\Gamma(\alpha)\beta^\alpha}\theta^{\alpha-1}e^{-\theta/\beta} = \frac{\theta^\alpha \tau x^{\tau-1}e^{-\theta(x^\tau+\beta^{-1})}}{\Gamma(\alpha)\beta^\alpha}.$$

From this we get

$$f(x) = \int_0^\infty \frac{\theta^\alpha \tau x^{\tau-1}e^{-\theta(x^\tau+\beta^{-1})}}{\Gamma(\alpha)\beta^\alpha} d\theta$$

$$= (x^\tau+\beta^{-1})^{-(\alpha+1)} \times \frac{\tau x^{\tau-1}\Gamma(\alpha+1)}{\beta^\alpha\Gamma(\alpha)}\int_0^\infty \frac{\theta^\alpha e^{-\theta(x^\tau+\beta^{-1})}}{\Gamma(\alpha+1)(x^\tau+\beta^{-1})^{-(\alpha+1)}} d\theta.$$

Since

$$\frac{\theta^\alpha e^{-\theta(x^\tau+\beta^{-1})}}{\Gamma(\alpha+1)(x^\tau+\beta^{-1})^{-(\alpha+1)}} \sim \Gamma\left(\alpha+1, \left(\frac{1}{\beta}(\beta x^\tau+1)\right)^{-1}\right),$$

We have that

$$\int_0^\infty \frac{\theta^\alpha e^{-\theta(x^\tau+\beta^{-1})}}{\Gamma(\alpha+1)(x^\tau+\beta^{-1})^{-(\alpha+1)}} d\theta = 1.$$

Therefore

$$f(x) = (x^\tau + \beta^{-1})^{-(\alpha+1)} \times \frac{\tau x^{\tau-1}\Gamma(\alpha+1)}{\beta^\alpha \Gamma(\alpha)} = \frac{\alpha \tau x^{\tau-1}}{\left(\frac{1}{\beta}\right)^{\alpha+1}(1+\beta x^\tau)^{\alpha+1}\beta^\alpha}$$

$$= \frac{\alpha\beta\tau x^{\tau-1}}{(1+\beta x^\tau)^{\alpha+1}}.$$

Therefore the compound (marginal) pdf of X is that of Burr.

4.1.5

a)

Since any continuous CDF has a Unif(0,1) distribution, $F(X_1) \sim Unif(0,1)$ and

$$P(X_1 \leq X_2) = E[P(X_1 \leq X_2|X_1)] = E[(1 - F_{X_2}(X_1))] = \int_0^1 u du = \frac{1}{2}.$$

Similarly

$$P(X_1 \leq X_i, i = 1, 2, ..., n) = E[P(X_1 \leq X_i, i = 1, 2, ..., n|X_1)]$$
$$= E[(1 - F(X_1))^{n-1}] = \int_0^1 u^{n-1} du = \frac{1}{n}.$$

b)

We have that

$$P(Y = j-1) = P(X_1 \leq X_2, ..., X_1 \leq X_{j-1}, X_1 > X_j) = E[(1 - F(X_1))^{j-2} F(X_1)]$$
$$= \int_0^1 u^{j-2}(1-u)du = \frac{1}{j(j-1)}.$$

Therefore

$$P(Y = y) = \frac{1}{y(y+1)}, y = 1, 2, 3, ...$$

c)

$$E(Y) = \sum_{y=1}^{\infty} \frac{y}{y(y+1)} = \sum_{y=1}^{\infty} \frac{1}{y+1} = \infty.$$

$$Var(Y) = \sum_{y=1}^{\infty} \frac{y}{y+1} - \left(\sum_{y=1}^{\infty} \frac{1}{1+y}\right)^2 = \infty$$

Therefore both the mean and variance of Y do not exist.

4.2.8

We want to find n so that

$$P(\overline{X} - \frac{1}{2} < \mu < \overline{X} + \frac{1}{2}) = 0.954$$

From this we get that $\alpha = 0.046$ and (using a table) $z_{\alpha/2} = 1.995$. The confidence interval is

$$(\overline{x} - z_{\alpha/2}\sigma/\sqrt{n}, \overline{x} + z_{\alpha/2}\sigma/\sqrt{n}).$$

From this we get

$$z_{\alpha/2}\sigma/\sqrt{n} = \frac{1}{2} \Rightarrow 1.995 \times \sqrt{10} \times 2 = \sqrt{n} \Rightarrow n = 159.201.$$

Therefore the value of n so that the probability is approximately 0.954 that the random interval $(\overline{X} - \frac{1}{2}, \overline{X} - \frac{1}{2})$ includes μ is 160.

4.2.18

a)

$$P(a < (n-1)S^2/\sigma^2 < b) = P\left(\frac{1}{b} < \frac{\sigma^2}{(n-1)S^2} < \frac{1}{a}\right)$$
$$= P\left(\frac{(n-1)S^2}{b} < \sigma^2 < \frac{(n-1)S^2}{a}\right) = 0.95$$

b)

Since $n = 9$ and $s^2 = 7.93$,

$$P\left(\frac{(8)(7.93)}{b} < \sigma^2 < \frac{(8)(7.93)}{a}\right) = 0.95,$$

where $b = \chi^2_{(8,0.975)} = 17.535$ and $a = \chi^2_{(8,0.025)} = 2.18$. Therefore the interval

$$\left(\frac{(8)(7.93)}{b}, \frac{(8)(7.93)}{a}\right) = \left(\frac{(8)(7.93)}{17.535}, \frac{(8)(7.93)}{2.18}\right) = (3.618, 29.101).$$

c)

If μ is known, we know that $\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2_{(n)}$. Then

$$P\left(\chi^2_{(n,a)} < \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 < \chi^2_{(n,b)}\right) = P\left(\frac{1}{\chi^2_{(n,b)}} < \frac{\sigma^2}{\sum_{i=1}^{n}(X_i-\mu)^2} < \frac{1}{\chi^2_{(n,a)}}\right)$$
$$= P\left(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{(n,b)}} < \sigma^2 < \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{(n,a)}}\right)$$

4.4.6

a)

Let m be the median of the distribution. Then

$$F(m) = P(X \leq m) = \int_0^m 2x dx = m^2 = 0.5.$$

$$\Rightarrow m = \sqrt{\frac{1}{2}}$$

Now

$$F(X_{(1)}) = P(min(X_1, X_2, X_3) \leq y)$$

$$= 1 - P((X_1, X_2, X_3) > y) = 1 - P(X_1 > y)P(X_2 > y)P(X_3 > y) = 1 - (1 - x^2)^3.$$

From this we get

$$P\left(X_{(1)} > \sqrt{\frac{1}{2}}\right) = 1 - P\left(X_{(1)} \leq \sqrt{\frac{1}{2}}\right) = 1 - \left(1 - \left(1 - \sqrt{\frac{1}{2}}^2\right)^3\right) = \frac{1}{8}.$$

b)

First, we find that

$$f(y_2) = 3!(y_2^2)(1 - y_2^2)^{3-2}(2y_2) = 12y_2^3(1 - y_2^2) \;\; for \;\; 0 < y_2 < 1$$

$$f(y_3) = 3(y_3^2)^{3-1}(1 - y_3^2)^{3-3}2y_3 = 6y_3^5 \;\; for \;\; 0 < y_3 < 1$$

$$f(y_2, y_3) = 3!(y_2^2)^{2-1}(y_3^2 - y_2^2)^0(1 - y_3^2)^0 2y_2 2y_3 = 24y_2^3 y_3 \;\; for \;\; 0 < y_2 \leq y_3 < 1$$

Now we get that

$$E(Y_2) = \int_0^1 12y_2^4(1 - y_2^2)dy_2 = \frac{24}{35}$$

$$E(Y_3) = \int_0^1 6y_3^6 dy_3 = \frac{6}{7}$$

$$Var(Y_2) = \frac{1}{2} - \left(\frac{24}{35}\right)^2$$

$$Var(Y_3) = \frac{3}{4} - \left(\frac{6}{7}\right)^2$$

$$E(Y_2 Y_3) = \int_0^1 \int_0^{y_3} 24y_2^4 y_3^2 dy_2 dy_3 = \frac{3}{5}$$

Therefore

$$\rho(Y_2, Y_3) = \frac{cov(Y_2, Y_3)}{\sqrt{Var(Y_2)Var(Y_3)}} = 0.5734$$

4.4.22

a)

We have that

$$Y_1 = \frac{Z_1}{n}$$

From this we get

$$Z_2 = (n-1)(Y_2 - \frac{Z_1}{n}) \Rightarrow Y_2 = \frac{Z_2}{n-1} + \frac{Z_1}{n}$$

Similarly,

$$Y_3 = \frac{Z_3}{n-2} + \frac{Z_2}{n-1} + \frac{Z_1}{n}, ..., Y_n = \sum_{i=1}^{n} \frac{Z_i}{n-i+1}$$

Therefore we get the jacobian

$$\mathbb{J} = \begin{vmatrix} \frac{1}{n} & 0 & 0 & 0 & \cdots & 0 \\ \frac{1}{n} & \frac{1}{n-1} & 0 & 0 & \cdots & 0 \\ \frac{1}{n} & \frac{1}{n-1} & \frac{1}{n-2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ \frac{1}{n} & & & & & 1 \end{vmatrix} = \frac{1}{n!}.$$

We also know that

$$f(y_1, y_2, ..., y_n) = n! f(y_1) \cdot \dots \cdot f(y_n) = n! e^{-\sum_{i=1}^{n} y_i} \quad for \;\; 0 < y_1 < ... < y_n < \infty.$$

Now we have

$$f(z_1, z_2, ..., z_n) = |J| f(y_1, y_2, ..., y_n) = \frac{n!}{n!} f(y_1) f(y_2) \cdot \dots \cdot f(y_n) \quad for \;\; 0 < y_1 < ... < y_n$$

$$= e^{-\frac{z_1}{n}} e^{-\frac{z_2}{n-1} - \frac{z_1}{n}} \cdot \dots \cdot e^{-\sum_{i=1}^{n} \frac{z_i}{n-i+1}} = e^{-\sum_{i=1}^{n} \frac{z_1}{n} - \sum_{i=1}^{n-1} \frac{z_2}{n-1} - \dots - z_n} = e^{-z_1 - z_2 - \dots - z_n} \quad for \;\; z_i > 0.$$

From this we can see that the $Z_i's$ are clearly independent since
$f(z_1, ..., z_n) = f(z_1) \cdot \dots \cdot f(z_n)$, and that each Z_i has the exponential distribution.

b)

$$\sum_{i=1}^{n} a_i Y_i = a_1 Y_1 + a_2 Y_2 + ... + a_n Y_n.$$

In part a, we have shown that $Z_1 = nY_1, Z_2 = (n-1)(Y_2 - Y_1), ..., Z_n = Y_n - Y_{n-1}$ are independent. We also had that

$$Y_1 = \frac{Z_1}{n}, Y_2 = \frac{Z_2}{n-1} + \frac{Z_1}{n}, ..., Y_n = Z_n + ... + \frac{Z_1}{n}.$$

Now

$$\sum_{i=1}^{n} a_i Y_i = a_1 \left(\frac{Z_1}{n} \right) + a_2 \left(\frac{Z_2}{n-1} + \frac{Z_1}{n} \right) + ... + a_n \left(Z_n + ... + \frac{Z_1}{n} \right).$$

Therefore we can see that all $Y_1, ..., Y_n$ can be expressed as linear functions of independent random variables, since in part a we have already shown that $Z_1, ..., Z_n$ are independent.

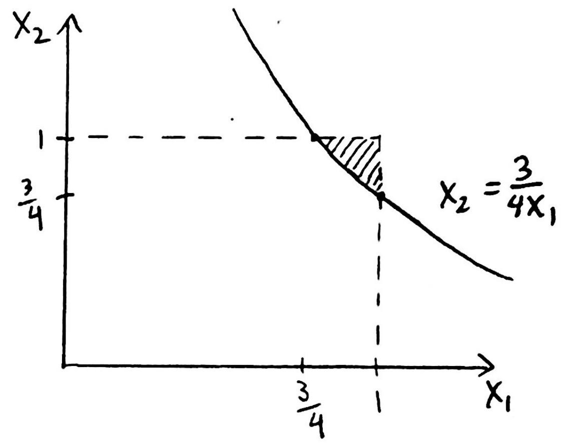
4.5.3

We have that

$$\gamma(\theta) = P_\theta((X_1, X_2) \in C) = P_\theta \left(X_1 X_2 \geq \frac{3}{4} \right).$$

Now

$$\because X_1 X_2 \geq \frac{3}{4} \Rightarrow X_2 \geq \frac{3}{4X_1}$$



So we get

$$\gamma(\theta) = \int_{\frac{3}{4}}^{1} \int_{\frac{3}{4x_1}}^{1} \theta^2 (x_1 x_2)^{\theta-1} dx_2 dx_1 = \int_{\frac{3}{4}}^{1} \theta x_1^{\theta-1} - \frac{\theta}{x_1}\left(\frac{3}{4}\right)^{\theta} dx_1 =$$

$$= 1 - \left(\frac{3}{4}\right)^{\theta} + \theta \left(\frac{3}{4}\right)^{\theta} log\left(\frac{3}{4}\right) \quad for \;\; \theta = 1, 2.$$

$\gamma(1)$ is the significance, and $\gamma(2)$ is the power when $\theta = 2$.