P8104 Homework Assignment 10

Yongyan Liu (yl6107)

Due Wed 11/26, 11:59pm

Problem 1

Let X and Y have a bivariate normal distribution with parameters $\mu_1 = 5$, $\mu_2 = 10$, $\sigma_1^2 = 1$, $\sigma_2^2 = 25$, and $\rho > 0$. If $P(4 < Y < 16|X = 5) = 0.954$, determine ρ.

Solution:

For a bivariate normal distribution, the conditional distribution of Y given $X = x$ is:

$$Y|X = x \sim N\left(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1), \sigma_2^2(1 - \rho^2)\right)$$

Given the parameters: - $\mu_1 = 5$, $\mu_2 = 10$ - $\sigma_1^2 = 1 \Rightarrow \sigma_1 = 1$ - $\sigma_2^2 = 25 \Rightarrow \sigma_2 = 5$

When $X = 5$:

$$\mu_{Y|X=5} = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(5 - \mu_1) = 10 + \rho \cdot \frac{5}{1} \cdot (5 - 5) = 10$$

$$\sigma_{Y|X=5}^2 = \sigma_2^2(1 - \rho^2) = 25(1 - \rho^2)$$

So $Y|X = 5 \sim N(10, 25(1 - \rho^2))$.

Now we use the given probability:

$$P(4 < Y < 16|X = 5) = 0.954$$

Standardizing:

$$P\left(\frac{4 - 10}{5\sqrt{1 - \rho^2}} < Z < \frac{16 - 10}{5\sqrt{1 - \rho^2}}\right) = 0.954$$

Let $z^* = \frac{6}{5\sqrt{1-\rho^2}}$. Then:

$$P(-z^* < Z < z^*) = 2\Phi(z^*) - 1 = 0.954$$

$$\Phi(z^*) = \frac{0.954 + 1}{2} = 0.977$$

From the standard normal table, $\Phi(2) = 0.9772 \approx 0.977$, so $z^* = 2$.

Therefore:

$$\frac{6}{5\sqrt{1 - \rho^2}} = 2$$

$$\rho = 0.8 \quad (\text{since } \rho > 0)$$

Problem 2

Let $X \sim N(0,1)$, $Y = X^2$

(a) What is the distribution of Y?

By definition, if $X \sim N(0,1)$, then $Y = X^2 \sim \chi^2_{(1)}$ (chi-squared distribution with 1 degree of freedom).

(b) Calculate the mean and variance of Y.

Solution:

Mean of Y:

$$E[Y] = E[X^2] = \text{Var}(X) + (E[X])^2 = 1 + 0 = 1$$

Variance of Y using Stein's Lemma:

Stein's Lemma: When $X \sim N(\mu, \sigma^2)$, $E[g(X)(X-\mu)] = \sigma^2 E[g'(X)]$.

For $X \sim N(0,1)$, we have $\mu = 0$ and $\sigma^2 = 1$: $E[g(X) \cdot X] = E[g'(X)]$

Therefore:

$$E[X^4] = E[3X^2] = 3E[X^2] = 3 \cdot 1 = 3$$
$$\text{Var}(Y) = E[Y^2] - (E[Y])^2 = E[X^4] - (E[X^2])^2 = 3 - 1 = 2$$

(c) If $W_i \sim \chi^2_{(1)}$, with all W_i independent. Define $U = \sum_{i=1}^r W_i$. What is the distribution of U?

Solution:

By the reproductive property of chi-squared distributions:

If $W_1, W_2, \ldots, W_r$ are independent with $W_i \sim \chi^2_{(1)}$, then:

$$U = \sum_{i=1}^{r} W_i \sim \chi^2_{(r)}$$

Proof using MGFs:

The MGF of $\chi^2_{(1)}$ is $M_{W_i}(t) = (1-2t)^{-1/2}$ for $t < 1/2$.

By independence:

$$M_U(t) = \prod_{i=1}^{r} M_{W_i}(t) = \left[(1-2t)^{-1/2}\right]^r = (1-2t)^{-r/2}$$

This is the MGF of $\chi^2_{(r)}$.

(d) Calculate the mean and variance of U.

Solution:

Since $U = \sum_{i=1}^r W_i$ where $W_i \sim \chi^2_{(1)}$ are independent:

Mean:

$$E[U] = \sum_{i=1}^{r} E[W_i] = \sum_{i=1}^{r} 1 = r$$

Variance:

$$\text{Var}(U) = \sum_{i=1}^{r} \text{Var}(W_i) = \sum_{i=1}^{r} 2 = 2r$$

Problem 3

Let $U \sim \chi^2_p$ and $V \sim \chi^2_q$, where U and V are independent of each other. Define

$$X = \frac{U/p}{V/q}$$

(a) Show that $X \sim F_{p,q}$.

Solution:

This is the **definition** of the F-distribution.

By definition, if $U \sim \chi^2_p$ and $V \sim \chi^2_q$ are independent, then:

$$X = \frac{U/p}{V/q} \sim F_{p,q}$$

(b) Derive the mean and variance of X.

Solution:

Mean of X:

$$E[X] = E\left[\frac{U/p}{V/q}\right] = \frac{q}{p}E[U] \cdot E\left[\frac{1}{V}\right]$$

Since $U \sim \chi^2_p$: $E[U] = p$

Since $V \sim \chi^2_q = \text{Gamma}(q/2, 1/2)$ (rate parameterization), we use the formula for negative moments of Gamma distribution: For $X \sim \text{Gamma}(\alpha, \beta)$, $E[X^{-k}] = \frac{\beta^k}{(\alpha-1)(\alpha-2)\cdots(\alpha-k)}$ for $\alpha > k$.

$$E\left[\frac{1}{V}\right] = E[V^{-1}] = \frac{1/2}{q/2 - 1} = \frac{1}{q-2}$$

(This requires $q > 2$.)

Therefore:

$$E[X] = \frac{q}{p} \cdot p \cdot \frac{1}{q-2} = \frac{q}{q-2}, \quad q > 2$$

Variance of X:

For $q > 4$:

$$E[X^2] = \frac{q^2}{p^2}E[U^2] \cdot E\left[\frac{1}{V^2}\right]$$

$E[U^2] = \text{Var}(U) + (E[U])^2 = 2p + p^2 = p(p+2)$

$E[V^{-2}] = \frac{(1/2)^2}{(q/2-1)(q/2-2)} = \frac{1/4}{(q-2)(q-4)/4} = \frac{1}{(q-2)(q-4)}$ (requires $q > 4$)

$$E[X^2] = \frac{q^2}{p^2} \cdot p(p+2) \cdot \frac{1}{(q-2)(q-4)} = \frac{q^2(p+2)}{p(q-2)(q-4)}$$

$$\text{Var}(X) = E[X^2] - (E[X])^2 = \frac{q^2(p+2)}{p(q-2)(q-4)} - \frac{q^2}{(q-2)^2}$$

(c) Show that $1/X \sim F_{q,p}$.

Solution:

$$\frac{1}{X} = \frac{V/q}{U/p}$$

Since $V \sim \chi_q^2$ and $U \sim \chi_p^2$ are independent, by the definition of F-distribution:

$$\frac{1}{X} = \frac{V/q}{U/p} \sim F_{q,p}$$

(d) Show that the median of an $F_{p,p}$ random variable equals 1 for any p.

Solution:

Let $X \sim F_{p,p}$. We need to show that $P(X \leq 1) = 0.5$.

From part (c), if $X \sim F_{p,p}$, then $1/X \sim F_{p,p}$ as well. This means X and $1/X$ have the same distribution.

For any $c > 0$:

$$P(X \leq c) = P\left(\frac{1}{X} \geq \frac{1}{c}\right) = 1 - P\left(\frac{1}{X} < \frac{1}{c}\right)$$

Since $1/X \stackrel{d}{=} X$:

$$P(X \leq c) = 1 - P\left(X < \frac{1}{c}\right)$$

Setting $c = 1$ and sicne X is continuous:

$$P(X \leq 1) = 1 - P(X \leq 1)$$

$$P(X \leq 1) = 0.5$$

Therefore, the median of $F_{p,p}$ is 1.

(e) Show that $\dfrac{(p/q)X}{1 + (p/q)X} \sim \text{Beta}\left(\frac{p}{2}, \frac{q}{2}\right)$.

Solution:

Let $W = \frac{(p/q)X}{1+(p/q)X}$.

Since $X = \dfrac{U/p}{V/q}$, we have $\dfrac{p}{q}X = \dfrac{U}{V}$.

Therefore:

$$W = \frac{U/V}{1 + U/V} = \frac{U}{V + U} = \frac{U}{U + V}$$

Since $U \sim \chi_p^2 = \text{Gamma}(p/2, 1/2)$ and $V \sim \chi_q^2 = \text{Gamma}(q/2, 1/2)$ are independent with the **same scale parameter**, the ratio:

$$W = \frac{U}{U + V} \sim \text{Beta}\left(\frac{p}{2}, \frac{q}{2}\right)$$

This follows from the well-known property: If $U \sim \text{Gamma}(\alpha, \lambda)$ and $V \sim \text{Gamma}(\beta, \lambda)$ are independent, then $\frac{U}{U+V} \sim \text{Beta}(\alpha, \beta)$.

Problem 4

Let $X_1, \ldots, X_n$ be a random sample from a population with pdf

$$f_X(x) = \begin{cases} \frac{1}{\theta}, & 0 < x < \theta \\ 0, & \text{otherwise} \end{cases}$$

Let $X_{(1)}, \ldots, X_{(n)}$ be the order statistics. Show that $\frac{X_{(1)}}{X_{(n)}}$ and $X_{(n)}$ are independent.

Solution:

First, we find the joint pdf of $(X_{(1)}, X_{(n)})$.

The joint pdf of the minimum and maximum order statistics is:

$$f_{X_{(1)},X_{(n)}}(u, v) = n(n-1)[F(v) - F(u)]^{n-2} f(u)f(v)$$

for $0 < u < v < \theta$.

For Uniform$(0, \theta)$: $F(x) = x/\theta$ and $f(x) = 1/\theta$.

$$f_{X_{(1)},X_{(n)}}(u, v) = n(n-1)\left(\frac{v-u}{\theta}\right)^{n-2} \cdot \frac{1}{\theta} \cdot \frac{1}{\theta}$$

$$= \frac{n(n-1)}{\theta^n}(v-u)^{n-2}, \quad 0 < u < v < \theta$$

Let $R = \frac{X_{(1)}}{X_{(n)}}$ and $S = X_{(n)}$. Then $X_{(1)} = RS$ and $X_{(n)} = S$.

The Jacobian is:

$$J = \begin{vmatrix} \frac{\partial u}{\partial r} & \frac{\partial u}{\partial s} \\ \frac{\partial v}{\partial r} & \frac{\partial v}{\partial s} \end{vmatrix} = \begin{vmatrix} s & r \\ 0 & 1 \end{vmatrix} = s$$

The joint pdf of (R, S) is:

$$f_{R,S}(r, s) = f_{X_{(1)},X_{(n)}}(rs, s) \cdot |J| = \frac{n(n-1)}{\theta^n} s^{n-1}(1-r)^{n-2}$$

The support is: $0 < rs < s < \theta$, which gives $0 < r < 1$ and $0 < s < \theta$.

This can be factored as:

$$f_{R,S}(r, s) = \underbrace{(n-1)(1-r)^{n-2}}_{g(r)} \cdot \underbrace{\frac{n \cdot s^{n-1}}{\theta^n}}_{h(s)}$$

for $0 < r < 1$ and $0 < s < \theta$.

Since the joint pdf factors into a function of r only times a function of s only, and the support is a product space, $R = \frac{X_{(1)}}{X_{(n)}}$ **and** $S = X_{(n)}$ **are independent**.

Problem 5

Let $X_1, \ldots, X_n$ be i.i.d lognormal(μ, σ^2) random variables. Let $G_n = \prod_{i=1}^n X_i^{1/n}$ be the sample Geometric mean.

(a) What is the distribution of G_n?

Solution:

If $X_i \sim \text{lognormal}(\mu, \sigma^2)$, then $\ln X_i \sim N(\mu, \sigma^2)$.

Taking the logarithm of G_n:

$$\ln G_n = \ln\left(\prod_{i=1}^n X_i^{1/n}\right) = \frac{1}{n}\sum_{i=1}^n \ln X_i$$

Let $Y_i = \ln X_i \sim N(\mu, \sigma^2)$. Then:

$$\ln G_n = \frac{1}{n}\sum_{i=1}^n Y_i = \bar{Y}$$

Since $Y_1, \ldots, Y_n$ are i.i.d. $N(\mu, \sigma^2)$:

$$\bar{Y} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Therefore, $\ln G_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, which means:

$$G_n \sim \text{lognormal}\left(\mu, \frac{\sigma^2}{n}\right)$$

(b) Show that $G_n \xrightarrow{P} e^\mu$.

Solution:

We need to show that G_n converges in probability to e^μ.

From part (a), $\ln G_n = \bar{Y}$ where $Y_i = \ln X_i \sim N(\mu, \sigma^2)$ are i.i.d.

By the **Weak Law of Large Numbers (WLLN)**:

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^n Y_i \xrightarrow{P} E[Y_1] = \mu$$

Since the exponential function is continuous:

$$G_n = e^{\ln G_n} = e^{\bar{Y}} \xrightarrow{P} e^\mu$$