IE 605: Engineering Statistics

Solutions of tutorial 5

Solution 1

Part 1: Add and subtract $\overline{x}$ from the expression on the left hand side and then expand as follows

$$\sum_{i=1}^{n}(x_i - \overline{x} + \overline{x} - a)^2 = \sum_{i=1}^{n}(x_i - \overline{x})^2 + 2\sum_{i=1}^{n}(x_i - \overline{x})(\overline{x} - a) + \sum_{i=1}^{n}(\overline{x} - a)^2$$

we can write the middle term in above eqn. ans simplify

$$\begin{aligned}
\sum_{i=1}^{n}(x_i - \overline{x})(\overline{x} - a) &= \overline{x}\sum_{i=1}^{n} x_i - a\sum_{i=1}^{n} x_i - \overline{x}\sum_{i=1}^{n}\overline{x} + \overline{x}\sum_{i=1}^{n} a \\
&= n\overline{x}^2 - an\overline{x} - n\overline{x}^2 + n\overline{x}a \\
&= 0
\end{aligned}$$

Then, we get

$$\sum_{i=1}^{n}(x_i - a)^2 = \sum_{i=1}^{n}(x_i - \overline{x})^2 + \sum_{i=1}^{n}(\overline{x} - a)^2 \tag{1}$$

Now, it can be verifies that above eqn. attains minimum at $a = \overline{x}$.

Part 2: Now, consider

$$\begin{aligned}
s^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \\
&= \frac{1}{n-1}\left\{\sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\bar{x}^2\right\} \\
&= \frac{1}{n-1}\left\{\sum_{i=1}^{n} x_i^2 - 2n\bar{x}^2 + n\bar{x}^2\right\} \\
&= \frac{1}{n-1}\left\{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right\}
\end{aligned}$$

Solution 2

Part 1

The pdf of $V \sim \chi^2_{n-1}$ is given by

$$f_V(v) = \frac{1}{2^{(\frac{n-1}{2})}\Gamma(\frac{n-1}{2})} v^{\frac{n-1}{2}-1} e^{-v/2}; \quad v > 0$$

The pdf of $U \sim \mathcal{N}(0,1)$ is given by

$$f_U(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}; \quad u \in \mathbb{R}$$

Since U and V are independent , we can write their joint pdf as:

$$\begin{aligned}
f_{UV}(u,v) &= f_U(u) f_V(v) \\
&= \frac{1}{\sqrt{2\pi} 2^{(\frac{n-1}{2})}\Gamma(\frac{n-1}{2})} e^{-u^2/2} v^{\frac{n-1}{2}-1} e^{-v/2}, \quad u \in \mathbb{R}, v > 0 \\
&= \frac{1}{2^{(\frac{n}{2})}\Gamma(\frac{n-1}{2})\Gamma(\frac{1}{2})} e^{-u^2/2} v^{\frac{n-3}{2}} e^{-v/2} \quad \text{since, } \Gamma(\frac{1}{2}) = \sqrt{\pi}
\end{aligned}$$

Part 2:

Given, $X = \frac{U}{\sqrt{\frac{V}{n-1}}}$ and $Y = V$ Then the inverses are given as, $U = X\sqrt{\frac{Y}{n-1}}$ and $V = Y$ and the jacobian is given as,

$$\mathbb{J} = \begin{bmatrix} \sqrt{\frac{Y}{n-1}} & \frac{1}{2(n-1)\sqrt{\frac{Y}{n-1}}} \\ 0 & 1 \end{bmatrix} = \sqrt{\frac{Y}{n-1}}$$

Now, the joint pdf of X and Y is given as,

$$\begin{aligned}
f_{XY}(x,y) &= f_{UV}(u,v)\mathbb{J} \quad \text{,where } u = x\sqrt{\frac{y}{n-1}} \text{ and } v = y \\
&= \frac{1}{2^{(\frac{n}{2})}\Gamma(\frac{n-1}{2})\Gamma(\frac{1}{2})} e^{-(x\sqrt{\frac{y}{n-1}})^2/2} y^{\frac{n-3}{2}} e^{-y/2} \sqrt{\frac{y}{n-1}} \\
&= \frac{1}{2^{(\frac{n}{2})}\Gamma(\frac{n-1}{2})\Gamma(\frac{1}{2})\sqrt{n-1}} e^{-y(\frac{x^2}{2(n-1)}+\frac{1}{2})} y^{\frac{n}{2}-1} \quad \text{where } y > 0 \text{ and } x \in \mathbb{R}
\end{aligned}$$

Part 3

The marginal distribution of X is given by,

$$f_X(x) = \int_0^\infty f_{XY}(x,y)\, dy$$

$$
\begin{aligned}
&= \int_0^\infty \frac{1}{2^{(\frac{n}{2})}\Gamma(\frac{n-1}{2})\Gamma(\frac{1}{2})\sqrt{n-1}} e^{-y(\frac{x^2}{2(n-1)}+\frac{1}{2})} y^{\frac{n}{2}-1} dy \\
&= \frac{1}{2^{(\frac{n}{2})}\Gamma(\frac{n-1}{2})\Gamma(\frac{1}{2})\sqrt{n-1}} \int_0^\infty e^{-y(\frac{x^2}{2(n-1)}+\frac{1}{2})} y^{\frac{n}{2}-1} dy \\
&= \frac{1}{2^{(\frac{n}{2})}\Gamma(\frac{n-1}{2})\Gamma(\frac{1}{2})\sqrt{n-1}} \Big(\frac{2(n-1)}{x^2+n-1}\Big)^{\frac{n}{2}} \int_0^\infty e^{-u} u^{\frac{n}{2}-1} du \\
&= \frac{1}{2^{(\frac{n}{2})}\Gamma(\frac{n-1}{2})\Gamma(\frac{1}{2})\sqrt{n-1}} \Big(\frac{2(n-1)}{x^2+n-1}\Big)^{\frac{n}{2}} \Gamma(\frac{n}{2}) \\
&= \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})\Gamma(\frac{1}{2})\sqrt{n-1}} \Big(\frac{1}{\frac{x^2}{n-1}+1}\Big)^{\frac{n}{2}} \\
\implies f_X(x) &= \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})\Gamma(\frac{1}{2})\sqrt{n-1}} \Big(\frac{1}{\frac{x^2}{n-1}+1}\Big)^{\frac{n}{2}}, \quad x \in \mathbb{R}
\end{aligned}
$$

Part 4

We know that t-distribution with n-1 degrees of freedom is given as,

$$
\implies f_T(t) = \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})\Gamma(\frac{1}{2})\sqrt{n-1}} \Big(\frac{1}{\frac{t^2}{n-1}+1}\Big)^{\frac{n}{2}}, \quad t \in \mathbb{R}
$$

The above distribution is same as the distribution of X obtained in **Part 3**. Hence, X has t-distribution with n-1 degrees of freedom

Solution 3

Let $X =$ number of defective parts in the sample. Then $X \sim$ hypergeometric$(N = 100, M, K)$ where $M =$ number of defectives in the lot and $K =$ sample size.

1. If there are 6 or more defectives in the lot, then the probability that the lot is accepted $(X = 0)$ is at most

$$
\begin{aligned}
P(X = 0|M = 100, N = 6, K) &= \frac{\binom{6}{0}\binom{94}{K}}{\binom{100}{K}} \\
&= \frac{(100-K)(100-K-5)}{100.\ldots.95}.
\end{aligned}
$$

By trial and error we find $P(X = 0) = .10056$ for $K = 31$ and $P(X = 0) = .09182$ for $K = 32$. So the sample size must be at least 32.

Solution 4

Calculating the cdf of $Z^2 = [\min(X,Y)]^2 \Rightarrow Z^2 > 0$, we obtain

$$\begin{aligned}
F_{Z^2}(z) &= \mathbb{P}\left\{(\min(X,Y))^2 \le z\right\} \\
&= \mathbb{P}\left\{-\sqrt{z} \le \min(X,Y) \le \sqrt{z}\right\} \\
&= \mathbb{P}\left\{\min(X,Y) \le \sqrt{z}\right\} - \mathbb{P}\left\{\min(X,Y) \le -\sqrt{z}\right\} \\
&= [1 - \mathbb{P}\left\{\min(X,Y) > \sqrt{z}\right\}] - [1 - \mathbb{P}\left\{\min(X,Y) > -\sqrt{z}\right\}] \\
&= \mathbb{P}\left\{\min(X,Y) > -\sqrt{z}\right\} - \mathbb{P}\left\{\min(X,Y) > \sqrt{z}\right\} \\
&= \mathbb{P}\left\{X > -\sqrt{z}\right\}\mathbb{P}\left\{Y > -\sqrt{z}\right\} - \mathbb{P}\left\{X > \sqrt{z}\right\}\mathbb{P}\left\{Y > \sqrt{z}\right\},
\end{aligned}$$

where we use the independence of X and Y. Since X and Y are identically distributed, $\mathbb{P}\{X > a\} = \mathbb{P}\{Y > a\} = 1 - F_X(a)$, so

$$\begin{aligned}
F_{Z^2}(z) &= (1 - F_X(-\sqrt{z}))^2 - (1 - F_X(\sqrt{z}))^2 \\
&= 1 - 2F_X(-\sqrt{z}),
\end{aligned}$$

since $1 - F_X(\sqrt{z}) = F_X(-\sqrt{z})$. Differentiating and substituting gives

$$\begin{aligned}
f_{Z^2}(z) &= \frac{d}{dz}F_{Z^2}(z) \\
&= f_X(-\sqrt{z})\frac{1}{\sqrt{z}} \\
&= \frac{1}{\sqrt{2\pi}}e^{-z/2}z^{-1/2}, \quad z > 0
\end{aligned}$$

the pdf of a χ^2_1 random variable.

Solution 5

It can be verified that $X^2 + Y^2 \sim \chi^2_2$. Thus

$$P(X^2 + Y^2 < 1) = \int_0^1 \frac{e^{-x/2}}{2}dx = 1 - \frac{1}{\sqrt{e}} = 0.3935.$$

Solution 6

The pdf of beta distribution with parameters α and β where both α and β are unknown is given by,

$$f(x|\alpha,\beta) = \frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}, 0 < x < 1, \alpha > 0, \beta > 0.$$

To show: Beta distribution with parameters α and β where both α and β are unknown belongs to Exponential family.

Exponential Family: A family of pdfs and pmfs is called an *Exponential family* if it can be expressed as

$$f(x|\theta) = h(x)c(\theta)exp\left(\sum_{i=1}^{k} w_i(\theta)t_i(x)\right),$$

where $h(x) \geq 0$ and $t_1(x), \ldots, t_k(x)$ are real-valued functions of the observation x (they cannot depend on θ), and $c(\theta) \geq 0$ and $w_1(\theta), \ldots, w_k(\theta)$ are real-valued functions of the possibly vector-valued parameter θ.

In this case:

$$f(x|\alpha, \beta) = I_{[0,1]}(x)\frac{1}{B(\alpha, \beta)} \exp\{(\alpha - 1)\log(x) + (\beta - 1)\log(1 - x)\}$$

where $h(x) = I_{[0,1]}(x), c(\alpha, \beta) = \frac{1}{B(\alpha,\beta)}, w_1(\alpha) = \alpha - 1, t_1(x) = \log(x), w_2(\beta) = \beta - 1, t_2(x) = \log(1 - x)$.

Solution 7

The sample mean is given by $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

The sample variance is given by $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

Here we use the theorem that states,

Theorem 1. *Let $X_1, X_2, \ldots, X_n$ are independent random variables. Let $g_i(X_i)$ be a function of only $x_i, i = 1, 2, \ldots, n$. Then the random variables $U_i = g_i(X_i), i = 1, 2, \ldots, n$ are mutually independent.*

Applying the above theorem, we can write S^2 as a function of $(n - 1)$ deviations.

$$\begin{aligned}
S^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\left((X_1 - \bar{X})^2 + \sum_{i=2}^{n}(X_i - \bar{X})^2\right) \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\left(\left(\sum_{i=2}^{n}(X_i - \bar{X})\right)^2 + \sum_{i=2}^{n}(X_i - \bar{X})^2\right) \quad \text{since, } \sum_{i=1}^{n}(X_i - \bar{X}) = 0
\end{aligned}$$

Thus S^2 can be written as a function only of $(X_2 - \bar{X}, \ldots, X_n - \bar{X})$. We will now show that these random variables are independent of $\bar{X}$. The joint pdf of the sample $X_1, X_2, \ldots, X_n$ is given by

$$f(x_1, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}}e^{-1/2\sum_{i=1}^{n} x_i}, \qquad -\infty < x_i < \infty, \forall i = 1, 2, \ldots, n$$

Make the transformations

$$y_1 = \bar{x},$$
$$y_2 = x_2 - \bar{x},$$
$$.$$
$$.$$
$$.$$
$$y_n = x_n - \bar{x},$$

This is a linear transformation with a Jacobian equal to $1/n$ (Verify it). We have

$$\begin{aligned} g(y_1, \ldots, y_n) &= \frac{n}{(2\pi)^{n/2}} e^{-(1/2)(y_1 - \sum_{i=2}^{n} y_i)^2} e^{-(1/2)\sum\limits_{i=2}^{n}(y_i + y_1)^2}, \quad -\infty < y_i < \infty, \forall i = 1, \ldots, n \\ &= \left\{ (\frac{n}{2\pi})^2 e^{(-ny_1^2)/2} \right\} \left\{ \frac{n^{1/2}}{(2\pi)^{(n-1)/2}} e^{-(1/2)(\sum\limits_{i=2}^{n} y_i^2 + (\sum\limits_{i=2}^{n} y_i)^2)} \right\}, \quad -\infty < y_i < \infty, \forall i = 1, .., n \end{aligned}$$

Since the joint pdf of $Y_1, \ldots, Y_n$ factors, it follows from theorem that Y_1 is independent of $Y_2, \ldots, Y_n$ and, hence, it follows from theorem that $\bar{X}$ is independent of S^2.

Solution 8

The pdf of $X \sim Gamma(\alpha, \lambda)$ is given by

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}; \quad 0 < x < \infty, \alpha > 0, \lambda > 0$$

The mgf is given by

$$\begin{aligned} M_x(t) &= \int_0^\infty e^{tx} \frac{\lambda^\alpha}{\Gamma(\alpha)\lambda^\alpha} x^{\alpha-1} e^{-\lambda x} dx \\ &= \int_0^\infty \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-(\lambda - t)x} \\ &= \frac{\lambda^\alpha}{(\lambda - t)^\alpha} \int_0^\infty \frac{(\lambda - t)^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-(\lambda-t)x} dx \end{aligned}$$

The function in the last (underbraced) integral is a p.d.f. of gamma distribution $\Gamma(\alpha, \lambda - t)$ and, therefore, it integrates to 1. We get,

$$M_X(t) = \left(\frac{\lambda}{\lambda - t} \right)^\alpha$$

Now,

$$
\begin{aligned}
E[X^k] &= \int_0^\infty x^k \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} dx \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{(\alpha+k)-1} e^{-\lambda x} dx \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+k)}{\lambda^{\alpha+k}} \int_0^\infty \frac{\lambda^{\alpha+k}}{\Gamma(\alpha+k)} x^{(\alpha+k)-1} e^{-\lambda x} dx \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+k)}{\lambda^{\alpha+k}} .1 \\
&= \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\lambda^k} \\
&= \frac{(\alpha+k-1)\Gamma(\alpha+k-1)}{\Gamma(\alpha)\lambda^k} \\
&= \frac{(\alpha+k-1)(\alpha+k-2)...\alpha\Gamma(\alpha)}{\Gamma(\alpha)\lambda^k} \\
&= \frac{(\alpha+k-1)...\alpha}{\lambda^k}
\end{aligned}
$$

Therefore, first moment is

$$E[X] = \frac{\alpha}{\lambda}$$

and the second moment is

$$E[X^2] = \frac{(\alpha+1)\alpha}{\lambda^2}$$

Solution 9

The pdf of $X \sim \chi_r^2$ is given by

$$f(x) = \frac{1}{2^{r/2}\Gamma(r/2)} x^{r/2-1} e^{-x/2}; \quad x > 0$$

The MGF of X is calculated as follows:

$$
\begin{aligned}
M_X(t) &= \mathbb{E}(e^{tX}) \\
&= \int_0^\infty e^{tx} \cdot \frac{1}{2^{r/2}\Gamma(r/2)} x^{r/2-1} e^{-x/2} dx \\
&= \int_0^\infty \frac{1}{2^{r/2}\Gamma(r/2)} x^{r/2-1} e^{-(1/2-t)x} dx \\
&= (\tfrac{1}{2} - t)^{-r/2} \cdot \frac{1}{2^{r/2}\Gamma(r/2)} \int_0^\infty y^{r/2-1} e^{-y} dy \quad \text{putting } y = (\frac{1}{2} - t)x \\
&= (1-2t)^{-r/2} \cdot \frac{1}{\Gamma(r/2)} \int_0^\infty y^{r/2-1} e^{-y} dy \\
&= (1-2t)^{-r/2}
\end{aligned}
$$

[The second half of the above expression is a p.d.f. of Gamma distribution, therefore it integrates to 1].

We then have,

$$M_X(t) = \frac{1}{(1-2t)^{r/2}}$$

Now, let us define $Y = \sum_{i=1}^{n} X_i$ where $X_i \sim \chi^2_{n_i}$ for $i = 1, \ldots, n$.

Then, MGF of Y is given by,

$$\begin{aligned}
M_Y(t) &= \mathbb{E}(e^{tY}) \\
&= \mathbb{E}\left(e^{t(\sum_{i=1}^{n} X_i)}\right) \\
&= \prod_{i=1}^{n} \mathbb{E}(e^{tX_i}) \quad \text{since } X_i's \text{ are independent} \\
&= \prod_{i=1}^{n} \frac{1}{(1-2t)^{n_i/2}} \quad \text{since } X_i \sim \chi^2_{n_i} \\
&= \frac{1}{(1-2t)^{p/2}} \quad \text{where } p = \sum_{i=1}^{n} n_i
\end{aligned}$$

By uniqueness of MGF, we can then conclude that $Y = \sum_{i=1}^{n} X_i \sim \chi^2_p$ where $p = \sum_{i=1}^{n} n_i$

Solution 10

Let $X_1, \ldots, X_n$ be iid with pdf $f_X(x)$. Let $\bar{X}$ be the sample mean and $X = n\bar{X}$.

Therefore, $X = X_1 + \cdots + X_n$. Then $\bar{X} = (1/n)X$, a scale transformation. Therefore the pdf of $\bar{X}$ is

$$f_{\bar{X}}(x) = \tfrac{1}{1/n} f_X(\tfrac{x}{1/n}) = n f_X(nx).$$

Solution 11

Given: Let $\bar{X}_n$and S_n^2 are sample mean and sample variance of random sample $X_1, X_2, \ldots, X_n$. Let X_{n+1}^{th} sample is obtained.

- To show: $\bar{X}_{n+1} = \frac{X_{n+1} + n\bar{X}_n}{n+1}$.

$$\begin{aligned}
\bar{X}_{n+1} &= \frac{\sum\limits_{i=1}^{n+1} X_i}{n+1} \\
&= \frac{X_{n+1} + \sum\limits_{i=1}^{n} X_i}{n+1} \\
&= \frac{X_{n+1} + n\bar{X}_n}{n+1} \quad \text{where } \bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.
\end{aligned}$$

- To show: $nS_{n+1}^2 = (n-1)S_n^2 + \left(\frac{n}{n+1}\right)(X_{n+1} - \bar{X}_n)^2$.

$$
\begin{aligned}
nS_{n+1}^2 &= \frac{n}{(n+1)-1}\sum_{i=1}^{n+1}(X_i - \bar{X}_{n+1})^2 \\
&= \sum_{i=1}^{n+1}\left(X_i - \frac{X_{n+1} + n\bar{X}_n}{n+1}\right)^2 \qquad \text{use } \bar{X}_{n+1} = \frac{X_{n+1} + n\bar{X}_n}{n+1} \\
&= \sum_{i=1}^{n+1}\left(X_i - \frac{X_{n+1}}{n+1} + \frac{n\bar{X}_n}{n+1}\right)^2 \\
&= \sum_{i=1}^{n+1}\left\{\left(X_i - \bar{X}_n\right) - \left(\frac{X_{n+1}}{n+1} - \frac{\bar{X}_n}{n+1}\right)\right\}^2 \qquad \text{we adjust by } \pm\bar{X}_n \\
&= \sum_{i=1}^{n+1}\left\{\left(X_i - \bar{X}_n\right)^2 - 2\left(X_i - \bar{X}_n\right)\left(\frac{X_{n+1} - \bar{X}_n}{n+1}\right) + \frac{1}{(n+1)^2}\left(X_{n+1} - \bar{X}_n\right)^2\right\}
\end{aligned}
$$

since $\sum_{i=1}^{n}(X_i - \bar{X}_n) = 0$, we get

$$
\begin{aligned}
nS_{n+1}^2 &= \sum_{i=1}^{n}(X_i - \bar{X}_n)^2 + (X_{n+1} - \bar{X}_n)^2 - 2\frac{(X_{n+1} - \bar{X}_n)^2}{n+1} + \frac{n+1}{(n+1)^2}(X_{n+1} - \bar{X}_n)^2 \\
&= (n-1)S_n^2 + \frac{n}{n+1}(X_{n+1} - \bar{X}_n)^2.
\end{aligned}
$$