Solutions of the practice exam of ESE326 F2017

1. $f_{XY}(x,y) = cxy$

(1) In order for f_{XY} to be a valid density function, it has to satisfy the following condition

$$\iint_{-\infty}^{\infty} f_{XY}\,dxdy = 1$$

i.e.,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} cxy\,dxdy = 1 \Rightarrow \int_{0}^{1}\int_{y}^{1} cxy\,dxdy = 1 \Rightarrow \int_{0}^{1}\frac{1}{2}c(y-y^3)dy = 1$$

Solving this equation gives, $c = 8$. i.e., $f_{XY}(x,y) = 8xy$

(2)

$$f_X(x) = \int_0^x f_{XY}(x,y)\,dy = \int_0^x 8xy\,dy = 4x^3, 0 < x < 1$$

$$E[X] = \int_0^1 x f_X(x)\,dx = \int_0^1 4x^4\,dx = \frac{4}{5}$$

$$E[X^2] = \int_0^1 x^2 f_X(x)\,dx = \int_0^1 4x^5\,dx = \frac{2}{3}$$

(3)

$$f_Y(y) = \int_y^1 f_{XY}(x,y)\,dx = \int_y^1 8xy\,dx = 4y(1-y^2), \qquad 0 < y < 1$$

$$E[Y] = \int_0^1 y f_Y(y)\,dy = \int_0^1 4y^2(1-y^2)\,dy = \frac{8}{15}$$

$$E[Y^2] = \int_0^1 y^2 f_Y(y)\,dy = \int_0^1 4y^3(1-y^2)\,dy = \frac{1}{3}$$

(4)

$$f_{XY}(x,y) = 8xy, f_X(x) = 4x^3, f_Y(y) = 4y(1-y^2)$$

Obviously, we don't have $f_{XY}(x,y) = f_X(x)f_Y(y)$ for all $x\ and\ y$. Hence, $X\ and\ Y$ are not independent

(5) First, let's find out the conditional density of $X\ given\ Y = y$

$$f_{X|y} = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{8xy}{4y(1-y^2)} = \frac{2x}{1-y^2}$$

Then, the conditional mean of $X\ given\ Y = y$ can be found as:

$$\mu_{X|y} = \int_y^1 x f_{X|y}\, dx = \int_y^1 \frac{8x^2 y}{4y(1-y^2)} dx = \frac{2(1-y^3)}{3(1-y^2)}$$

The curve of regression is not linear.

2.

(1) since $\bar{X}_1$ and $\bar{X}_2$ are the sample means of X_1 and X_2, we have,

$$E[\bar{X}_1\,] = \mu_1;\; E[\bar{X}_2\,] = \mu_2;\; \bar{X}_1 \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right);\; \bar{X}_2 \sim N(\mu_2, \frac{\sigma_2^2}{n_2})$$

Hence,

$$E[\bar{X}_1 - \bar{X}_2] = E[\bar{X}_1\,] - E[\bar{X}_2\,] = \mu_1 - \mu_2$$

This proves that $\bar{X}_1 - \bar{X}_2$ is an unbiased estimator of $\mu_1 - \mu_2$.

(2) The moment generating function of $\bar{X}_1$ and $\bar{X}_2$ are:

$$m_{\bar{X}_1}(t) = e^{\left(\mu_1 t + \frac{t^2}{2}(\frac{\sigma_1^2}{n_1})\right)}; m_{\bar{X}_2}(t) = e^{\left(\mu_2 t + \frac{t^2}{2}(\frac{\sigma_2^2}{n_2})\right)}$$

The moment generating function of $\bar{X}_1 - \bar{X}_2$ can be found as:

$$m_{\bar{X}_1 - \bar{X}_2}(t) = m_{\bar{X}_1}(t) m_{\bar{X}_2}(-t) = e^{\left(\mu_1 t + \frac{t^2}{2}(\frac{\sigma_1^2}{n_1})\right)} e^{\left(-\mu_2 t + \frac{t^2}{2}(\frac{\sigma_2^2}{n_2})\right)} = e^{\left((\mu_1 - \mu_2)t + \frac{t^2}{2}(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})\right)}$$

This proves that $\bar{X}_1 - \bar{X}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$

(3) When $\sigma_1^2 = \sigma_2^2 = \sigma^2$, we have,

$$\frac{(n_1 + n_2 - 2)S_p^2}{\sigma^2} = \frac{(n_1 + n_2 - 2)}{\sigma^2} \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = \frac{(n_1 - 1)S_1^2}{\sigma^2} + \frac{(n_2 - 1)S_2^2}{\sigma^2}$$

We already know that $\frac{(n_1 - 1)S_1^2}{\sigma^2} \sim \mathrm{X}_{n_1 - 1}^2$ and $\frac{(n_2 - 1)S_2^2}{\sigma^2} \sim \mathrm{X}_{n_2 - 1}^2$. The moment generating function of these two random variables are

$$m_1(t) = (1 - 2t)^{-(n_1 - 1)/2}, m_2(t) = (1 - 2t)^{-(n_2 - 1)/2}$$

The moment generating function of $\frac{(n_1 + n_2 - 2)S_p^2}{\sigma^2}$ can be found as:

$$m(t) = m_1(t) m_2(t) = (1 - 2t)^{-(n_1 + n_2 - 2)/2}$$

This proves that $\frac{(n_1 + n_2 - 2)S_p^2}{\sigma^2} \sim \mathrm{X}_{n_{1+n_2} - 2}^2$

(4) When $\sigma_1^2 = \sigma_2^2 = \sigma^2$, we have

$$\bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right) \Rightarrow \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \sim Z$$

And

$$\frac{(n_1+n_2-2)S_p^2}{\sigma^2} \sim \mathrm{X}^2_{n_{1+n_2}-2}$$

Then, following the definition of a T random variable, we have,

$$\frac{\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}}}}{\sqrt{\frac{\frac{(n_1+n_2-2)S_p^2}{\sigma^2}}{n_1+n_2-2}}} \sim T_{n_1+n_2-2} \Rightarrow \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{S_p^2(\frac{1}{n_1}+\frac{1}{n_2})}} \sim T_{n_1+n_2-2}$$

3.

(1) The hypothesis: H_0: $\sigma_1^2 = \sigma_2^2$; H_1: $\sigma_1^2 \neq \sigma_2^2$

The test statistic $\frac{S_1^2}{S_2^2} \sim F_{10,8}$. Assuming H_0 is true, the observed value of the test statistic is: $\frac{S_1^2}{S_2^2} = 1.33$

Since $\frac{S_1^2}{S_2^2} = 1.33 > 1$, and this is a two-tailed test, we choose the p-value as $p_{value} = 2P\left[F_{10,8} \geq 1.33\right]$

From the F probability table, we have, $P\left[F_{10,8} \leq 2.538\right] = 0.9, P\left[F_{10,8} \geq 2.538\right] = 0.1$

Hence, we have, $\left[F_{10,8} \geq 1.33\right] > 0.1$. i.e., $p_{value} > 0.2$

At the significance level of 0.2, H_0 will not be rejected. i.e., $\sigma_1^2 = \sigma_2^2$ still holds.

(2) The 95% confidence interval on $\mu_1 - \mu_2$ is: $(\bar{X}_1 - \bar{X}_2) \pm t_{0.025}\sqrt{{S_p}^2(\frac{1}{n_1}+\frac{1}{n_2})}$

$$S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2} = \frac{10\times10.27+8\times7.75}{18} = 9.15$$

Since the test statistic

$$\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{S_p^2(\frac{1}{n_1}+\frac{1}{n_2})}} \sim T_{18}$$

From the T-table, we find that $t_{0.025} = 2.101$

The 95% confidence interval is found as: 3.88 ± 2.86. Since this interval are positive throughout. We can say that with 95% confidence, $\mu_1 > \mu_2$

(3) Let's test on $H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$

Assuming H_0 is true, the test statistic

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2(\frac{1}{n_1} + \frac{1}{n_2})}} \sim T_{18}$$

The observed value of the test statistic is:

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_0}{\sqrt{S_p^2(\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{3.88}{1.36} = 2.85$$

Since this is a two-tailed test and 2.85>1, the p-value is chosen as:

$$p_{value} = 2P[T_{18} \geq 2.85]$$

From the T probability table, we find that,

$$P[T_{18} \geq 2.552] = 0.01, P[T_{18} \geq 2.878] = 0.005$$

Hence,

$$0.005 < P[T_{18} \geq 2.85] < 0.01$$

$$p_{value} < 0.02$$

Therefore, at the significance level of 0.1, H_0 will be rejected. i.e., $\mu_1 \neq \mu_2$

(4) Assuming H_0 is true, the test statistic:

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_0}{\sqrt{S_p^2(\frac{1}{n_1} + \frac{1}{n_2})}} \sim T_{18}$$

This is a two-tailed test. The rejection region consists of two parts $(-\infty, -t_0]$ and $[t_0, \infty)$. Because of the symmetry of the T random variable, the critical values are $\pm t_0$

t_0 can be found so that $P[type\ I\ error] = \alpha = 0.1$. i.e.,

$$P[T_{18} \geq t_0] + P[T_{18} \leq -t_0] = 0.1 \Rightarrow P[T_{18} \geq t_0] = 0.05 \Rightarrow t_0 = 1.734$$

Hence, the rejection region of the test is: $(-\infty, -1.734], and\ [1.734, \infty)$

Assuming H_0 is true, the observed value of the test statistic is 2.85 that is in the rejection region.

Therefore, at the significance level of 0.1, H_0 is rejected. i.e., $\mu_1 \neq \mu_2$