CENTRE FOR MATHEMATICAL SCIENCES
EXAMINER ANSWER SCRIPT (FINAL EXAM)
SEMESTER: I SESSION: 2019/2020

Course: APPLIED STATISTICS
Course Code: BUM2413

QUESTION 1

a

(i) The number of full processed bottles from the machine **or**
The number of not full processed bottles from the machine

(ii) $p_{full} = \dfrac{280}{300} = 0.9333$

Step 1:
$H_0 : \pi_{full} \geq 0.95$ ($Claim$: The machine is operating properly)
$H_1 : \pi_{full} < 0.95$

Step 2: $Z_{test} = \dfrac{0.9333 - 0.95}{\sqrt{\dfrac{0.95(0.05)}{300}}} = -1.3272$

Step 3: $-Z_{0.1} = -1.2816$

Step 4: Since $(Z_{test} = -1.3272) < (-Z_{0.1} = -1.2816)$, then reject H_0.

Step 5: At $\alpha = 0.1$, there is sufficient evidence to reject the claim. The machine is not operating properly based on data from the daily-checking.

OR $p_{Notfull} = \dfrac{20}{300} = 0..0667$

$H_0 : \pi_{Notfull} \geq 0.05$
$H_1 : \pi_{Notfull} < 0.05$ ($Claim$: The machine is operating properly)

Step 2: $Z_{test} = \dfrac{0.0667 - 0.05}{\sqrt{\dfrac{0.05(0.95)}{300}}} = 1.3245$

Step 3: $-Z_{0.1} = -1.2816$

Step 4: Since $(Z_{test} = 1.3272) > (-Z_{0.1} = -1.2816)$, then do not reject H_0.

Step 5: At $\alpha = 0.1$, there is insufficient evidence to support the claim. The machine is not operating properly based on data from the daily-checking.

<table>
<tr><td>b</td><td>(i) rural : $p_A = 0.7, \quad p_B = 0.3$ urban : $p_A = 0.5, \quad p_B = 0.5$

Step 1: $H_0 : \pi_R - \pi_U = 0$
$H_1 : \pi_R - \pi_U \neq 0 \quad (Claim)$

Step 2: $Z_{0.015} = 2.1701$

A 97% CI for $\pi_R - \pi_U = \left((0.7 - 0.5) \pm 2.1701 \left(\sqrt{\frac{0.7(0.3)}{100} + \frac{0.5(0.5)}{100}} \right) \right)$

$= (0.2 \pm 0.1472) = (0.0528, 0.3472)$

Step 3: Since $\pi_0 = 0$ is not included in the interval, then reject H_0.

Step 4: At $\alpha = 0.03$, there is sufficient evidence to support the claim. the proportions of all demands for Product A are different between rural and urban areas.</td></tr>
<tr><td></td><td>(ii) Two-tailed test</td></tr>
<tr><td colspan="2" align="center">QUESTION 2</td></tr>
<tr><td>i</td><td>$\frac{\sigma_X^2}{\sigma_Y^2}$ or population variance or population standard deviation</td></tr>
<tr><td>ii</td><td>Step 1: $H_0 : \sigma_X^2 \geq \sigma_Y^2$
$H_1 : \sigma_X^2 < \sigma_Y^2 \quad$ (claim)

Step 2: $P -$ value $= 0.4627$

Step 3: Since (P - value $= 0.4627$) > ($\alpha = 0.07$), we do not reject H_0.

Step 4: At $\alpha = 0.07$, there is insufficient evidence to support the claim. The data do not support the claim that Bank X provides more efficient service time than Bank Y.

OR

Step 1: $H_0 : \sigma_Y^2 \leq \sigma_Y^2$
$H_1 : \sigma_Y^2 > \sigma_X^2 \quad$ (claim)

Step 2: $P -$ value $= 0.4627$

Step 3: Since (P - value $= 0.4627$) > ($\alpha = 0.07$), we do not reject H_0.

Step 4: At $\alpha = 0.07$, there is insufficient evidence to support the claim. The data do not support the claim that Bank X provides more efficient service time than Bank Y.</td></tr>
<tr><td>iii</td><td>Bank X</td></tr>
<tr><td>iv</td><td>Since do not reject H_0, H_0 is false (Claim in H_1), therefore Type II error.</td></tr>
<tr><td colspan="2" align="center">QUESTION 3</td></tr>
<tr><td>a</td><td>(i) 4</td></tr>
<tr><td></td><td>(ii) df “Between Groups” : 3
df “Within Groups” : 16</td></tr>
</table>

	SSE: 0.0310 x16 = 0.496 **or** 2.1095 - 1.6135= 0.496 MStrt: 1.6135 / 3 = 0.5378 Ftest = 0.5378 / 0.0310 = 17.3484
	(iii) $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ $H_1: \mu_i \neq \mu_j$ for at least one $i \neq j$ where $i, j = 1, 2, 3, 4$ $P - value = 2.7734 \times 10^{-5}$ Since $\left(P - value = 2.7734 \times 10^{-5}\right) < (\alpha = 0.05)$, reject H_0. At $\alpha = 0.05$, at least one of the population mean powers differ for different panels. **OR** $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ $H_1: \mu_i \neq \mu_j$ for at least one $i \neq j$ where $i, j = 1, 2, 3, 4$ $f_{test} = 17.3484$ Since $(f_{test} = 17.3484) > (f_{critical} = 3.2389)$, reject H_0. At $\alpha = 0.05$, at least one of the population mean powers differ for different panels.
b	H_0: There is no effect of the glass particle size H_1: There is an effect of the glass particle size (B1) $p - value = 0.0010$ (B1) $p - value < \alpha = 0.01$, reject H_0 (M1 A1) At $\alpha = 0.01$, there is an effect of the glass particle size on strength of the brick (A1) **OR** H_0: There is no effect of the glass particle size H_1: There is an effect of the glass particle size (B1) $Ftest = 15.4271$ (B1) $Ftest = 15.4271 > Fcrit = 8.2854$, reject H_0 (M1 A1) At $\alpha = 0.01$, there is an effect of the glass particle size on strength of the brick (A1)
	QUESTION 4
i	Number of students.

ii	$H_0 : \beta_1 = 0$ or the slope is zero (no linear relationship between x and y) $H_1 : \beta_1 \neq 0$ or the slope is not zero (there is linear relationship) $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{2183.8333}{2774.9167} = 0.7870$ $t_{test} = \frac{\hat{\beta}_1 - \beta_1}{se\left(\hat{\beta}_1\right)} = \frac{0.7870 - 0}{\sqrt{0.1120}} = 2.3513$ $\alpha = 0.06,\ t_{0.03,10} = 2.1202\ ,\ -t_{0.03,10} = -2.1202$ Since $t_{test} = 2.3513 > t_{0.03,10} = 2.1202$, Reject H_0. At $\alpha = 0.06$, there exist significant relationship between number of students registered for the course and the number of books that will be sold
QUESTION 5	
i	20 patients
ii	Coefficient of determination: Adjusted R^2: 0.9301 93.01% of the variation in the risk of having strokes can be predicted by the age, blood pressure and smoking habit.
iii	**Step 1:** H_0 : Neither of independent variables are related to dependent variable H_1 : At least one of independent variables are related to dependent variable **Step 2:** $P-value = 0.0000$ or 4.6712 x 10^{-10} **Step 3:** $(P-value = 0.0000) < (\alpha = 0.05)$, H_0 is rejected. **Step 4**: At $\alpha = 0.05$, there is sufficient evidence to conclude that at least one of independent variables are related to dependent variable. **OR** **Step 1:** H_0 : Neither of independent variables are related to dependent variable H_1 : At least one of independent variables are related to dependent variable **Step 2:** $f_{test} = 85.3037$ **Step 3:** $\left(f_{test} = 85.3037\right) > \left(f_{0.05;3,16} = 3.4105\right)$, H_0 is rejected. **Step 4**: At $\alpha = 0.05$, there is sufficient evidence that at least one of independent variables are related to dependent variable.
iv	Based on the *P*-value in the coefficients table from **Figure 1**, identify the significant predictors. $x_1 : \left(P-value = 0.1971\right) > \left(\alpha = 0.05\right)$, do not reject H_0 $x_2 : \left(P-value = 0.1625\right) > \left(\alpha = 0.05\right)$, do not reject H_0 $x_3 : \left(P-value = 0.0000\right) < \left(\alpha = 0.05\right)$, reject H_0 Hence, x_3 is the significant predictor.
v	**Single variable:** x_3

	Reason: Since its model $\hat{y} = -0.2127 + 1.6613x_3$ has the lowest significant *P*-value and the highest coefficient $r^2 = 0.9328$ **The best model**: The best predictor model is $\hat{y} = -0.2127 + 1.6613x_3$. **Reason**: This model has the lowest significant *P*-value and the highest coefficient of determination.
vi	$\hat{y} = -0.2127 + 1.6613x_3$ $x_3 = 20 : \hat{y} = -0.2127 + 1.6613(20) = 33.0133\%$

QUESTION 6

a	Combine the categories
b	(i) H_0: The proportions of plates in the four categories are 10%, 70%, 15% and 5%, respectively. (Claim) H_1: The proportions of plates in the four categories are not 10%, 70%, 15% and 5%, respectively.

Test Statistic:

x_i	O_i	P_i	$E_i = nP_i$	$\frac{(O_i - E_i)^2}{E_i}$
Premium	19	0.10	200(0.10) = 20	$\frac{(19-20)^2}{20} = 0.0500$
Conforming	133	0.70	200(0.70) = 140	0.3500
Downgraded	35	0.15	200(0.15) = 30	0.8333
Unacceptable	13	0.05	200(0.05) = 10	0.9000
				$\chi^2_{test} = 2.1333$

Step 3: $\chi^2_{0.005,3} = 12.8382$

Step 4: Since $\left(\chi^2_{test} = 2.1333\right) < \left(\chi^2_{0.005,3} = 12.8382\right)$, do not reject H_0

Step 5: At 0.5% significance level, the engineer's claim is true.

	(ii) Good fit or Yes

QUESTION 7

i	Number of failure / Configuration / Failure modes
ii	Ratio-level / Nominal / Nominal
iii	H_0: Configuration does not affect the mode of failure. H_1: Configuration does affect the mode of failure

	Mode X	Mode Y	Mode Z	$x_{i\cdot}$
Configuration A	20	44	17	81

Configuration B	4	17	7	28
$x._j$	24	61	24	$x..$= 109

O_{ij}	$E_{ij}=\dfrac{n_{i.}\times n_{.j}}{n_{..}}$	$\dfrac{\left(O_{ij}-E_{ij}\right)^2}{E_{ij}}$
$O_{11}=20$	$E_{11}=\dfrac{81\times 24}{109}=17.8349$	$\dfrac{\left(20-17.8349\right)^2}{17.8349}=0.2628$
$O_{12}=44$	$E_{12}=\dfrac{81\times 61}{109}=45.3303$	$\dfrac{\left(44-45.3303\right)^2}{45.3303}=0.0390$
$O_{13}=17$	$E_{13}=\dfrac{81\times 24}{109}=17.8349$	$\dfrac{\left(17-17.8349\right)^2}{17.8349}=0.0391$
$O_{21}=4$	$E_{21}=\dfrac{28\times 24}{109}=6.1651$	$\dfrac{\left(4-6.1651\right)^2}{6.1651}=0.7604$
$O_{22}=17$	$E_{22}=\dfrac{28\times 61}{109}=15.6697$	$\dfrac{\left(17-15.6697\right)^2}{15.6697}=0.1129$
$O_{23}=7$	$E_{23}=\dfrac{28\times 24}{109}=6.1651$	$\dfrac{\left(7-6.1651\right)^2}{6.1651}=0.1131$
$\chi^2_{test}=\sum_{i=1}^{r}\sum_{j=1}^{c}\dfrac{\left(O_{ij}-E_{ij}\right)^2}{E_{ij}}=1.3273$		

$$\chi^2_{critical}=\chi^2_{\alpha,(r-1)(c-1)}$$
$$=\chi^2_{0.025,(2-1)(3-1)}$$
$$=\chi^2_{0.025,2}$$
$$=7.3778$$

Since $\left(\chi^2_{test}=1.3273\right)<\left(\chi^2_{critical}=7.3778\right)$, then, we do not reject H_0.

Hence, configuration does not affect the mode of failure at α = 0.005