Marking scheme-Final Semester Exam of Probability and Statistics (40 Marks)

Question1.

a. State that the following statements are true or false **(5marks)**

i. The quota sampling technique is one of the probability sampling method.

ii. Then moment generating function of $f(x)$ is given by $M_X(t) = E\left[e^{tx}\right]$ with $t \in \mathbb{R}$, then for continuous case we have: $M_X(t) = E\left[e^{tx}\right] = \int_{\Omega} e^{-tx} f(x)\,dx$

iii. A given estimator $\widehat{\theta}$ is an efficiency estimator of statistic θ if $\widehat{\theta}$ is biased estimator

iv. The Confidence Interval on the mean when sample size $n < 30$ and the variance of population is unknown, is given as:

$$CI = \left[\overline{X} - t_{\frac{\alpha}{2},n-1} \times \frac{s}{\sqrt{n}}, \overline{X} + t_{\frac{\alpha}{2},n-1} \times \frac{s}{\sqrt{n}}\right]$$

v. The Weighted Aggregate Index is given by the following formula: $\dfrac{\sum(P_n \times W)}{\sum(P_0 \times W)}$

b. Two catalysts are being analysed to determine how they affect the mean yield of a chemical process. Specifically, Catalyst 1 is currently in use, but Catalyst 2 is acceptable. Since Catalyst 2 is cheaper, it should be adopted, providing it does not change the process yield. Through the following data, compute estimation of population standard deviation. ***(5 marks)***

Type of Catalyst	Sample Size	Sample Mean	Variance
Catalyst 1	$n_1 = 8$	$\overline{x}_1 = 92.255$	$s_1^2 = (2.39)^2$
Catalyst 2	$n_2 = 8$	$\overline{x}_2 = 92.733$	$s_2^2 = (2.98)^2$

Answer:

a.

i. F **/1 mark**
ii. F **/1 mark**
iii. F **/1 mark**
iv. T **/1 mark**
v. T **/1 mark**

b.

The estimate of σ^2 is $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$ **/2 marks**

$$s_p^2 = \frac{(8-1)(2.39)^2 + (8-1)(2.98)^2}{8+8-2}$$ **/1 mark**

$s_p^2 = 7.30$ **/1 mark**

The estimate of σ is $\sqrt{s_p^2} = \sqrt{7.30} = 2.70$ **/1 mark**

Question2. We need to estimate the hourly salary of the company's employees. Consider a sample of 100 employees given us the mean $\overline{X} = \$5.50$. The population standard deviation is $\sigma = \$1$. Find a confidence interval with a confidence level of 95% for the population mean hourly salary μ. We assume that the population is finite of size $N = 800$. ***(10 marks)***

Answer:

1. The sample mean is $\bar{X} = \$5.50$

2. The standard error is $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \times \sqrt{1 - \frac{n}{N}}$ **/1 mark**

$= \frac{1}{\sqrt{100}} \times \sqrt{1 - \frac{100}{800}} = \frac{1}{\sqrt{100}} \times \sqrt{\frac{7}{8}} = 0.0935$ **/2 marks**

3. The confidence level at 95%, is $P[LB \leq \mu \leq UB\,] = 0.95$ **/1 mark**

4. The confidence interval is calculated as:

$$CI = \left[\bar{X} - z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} \times \sqrt{1 - \frac{n}{N}}, \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \times \sqrt{1 - \frac{n}{N}}\right]$$ **/2 marks**

$\Leftrightarrow CI = [5.50 - 1.96 \times 0.0935, 5.50 + 1.96 \times 0.0935\,]$ **/2 marks**

The confidence interval nearest of parameter μ is

$\$5.317 \leq \mu \leq \$5.683$ **/2 marks**

Question3. A sample $n = 20, \overline{x} = 4.0, s = 0.83$ is taken from a normally distributed population that has a unknown μ and unknown σ. Test the hypothesis that if the mean $\mu \neq 3.6$ at $\alpha = 0.05$. ***(10 marks)***

Answer:

1. Parameter of interest: Mean μ.
2. Hypotheses:

   $H_0 : \mu = 3.6$ versus $H_1 : \mu \neq 3.6$.
3. Significance level:

   $\alpha = 0.05$.
4. As the population is normally distributed and σ is unknown, we use the t statistic.

We have $s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{0.83}{\sqrt{20}} = 0.1856$ **/2 marks**

The specific value of the test statistic is

$$t_0 = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{4.0 - 3.6}{0.1856} = 2.1552$$ **/3 marks**

5. Critical value:

$t_{\frac{\alpha}{2},n-1} = t_{\frac{0.05}{2},20-1} = t_{0.025,19} = 2.0930$ **/2 marks**

6. As $|2.1552| > 2.0930$, **/1 mark** we reject the null hypothesis $H_0 : \mu = 3.6$ at the 0.05 level of significance. **/2 marks**

Question 4.

a. The following data relate to a set of commodities used in a particular process. Calculate Fisher's price indices for 2013. ***(7 marks)***

Commodity	2012		2013	
	Price	Quantity	Price	Quantity
Meat	36	100	40	95
Fish	80	12	90	10
Butter	45	16	41	18
Milk	5	41	6	1200

b. Two different processes are used to produce a cylindrical tube to be used in water supply. The diameters of the tubes are normally distributed with unknown and unequal variances. Two random samples are drawn from each process, and the corresponding results are given in next table. Find the specific value of the test statistic t_0. ***(3 marks)***

Process	Sample Size	Sample Mean	Variance
Process 1	$n_1 = 18$	$\bar{x}_1 = 37$ mm	$s_1^2 = 11^2$
Process 2	$n_2 = 22$	$\bar{x}_2 = 44$ mm	$s_2^2 = 6^2$

Answer:

a.

Commodity	2013			
	Laspeyres		Paasche	
	$p_n q_0$	$p_0 q_0$	$q_n p_n$	$p_0 q_n$
Meat	4000	3600	3800	3420
Fish	1080	960	900	800

Butter	656	720	738	810
Milk	246	205	7200	6000
Total	**5982** **/1 mark**	**5485** **/1 mark**	**12638** **/1 mark**	**11030** **/1 mark**

$L_p = {}^{5982}/_{5485} = 109.1$, **/1 mark**

$P_p = {}^{12638}/_{11030} = 114.6$, **/1 mark**

$F_p = \sqrt{109.1 \times 114.6} = 111.8$ **/1 mark**

$$\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$$

b. $t_0 = \dfrac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}}$ **/1 mark**

$= \dfrac{(\bar{x}_1-\bar{x}_2)-\delta_0}{\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}} = \dfrac{(37-44)-\delta_0}{\sqrt{\dfrac{11^2}{18}+\dfrac{6^2}{22}}}$ **/1 mark**

$t_0 = -2.4212$ **/1 mark**

Success and God be with you!!!