CSE 5301: 002 – Fall 2018
Exam 2 - Variant α, Tuesday 11/06/2018

Name: SOLUTION

Student ID:

(Not providing this information: -10 Points)
(Name missing in Individual pages: -5 Points)

Name: ______________________

Total Exam Points: 100

Score

Question	Points	Max Points
1		10
2		10
3		20
4		20
5		15
6		25
7		10*
Total		**100 (110*)**

*: Optional Extra Credit Question

Name: ________________________

Question 1 – 10 points

A security engineer is aware that the number of attacks on a system is uniformly distributed. She observes the number of attacks the server is subjected two over 10 days. The numbers are 35, 45, 44, 22, 37, 32, 36, 54, 35, 31. Find the parameters of the distribution

For a Uniform distribution $U(a,b)$,

$$\mu_1 = \frac{a+b}{2} \qquad \mu_2 = \frac{(b-a)^2}{12}$$

are the 1st & 2nd population moments.

The 1st & 2nd sample moments are

$$m_1 = \bar{X} = \frac{35+45+44+22+37+32+36+54+35+31}{10}$$

$$= 37.1$$

$$m_2 = \bar{s}^2 = \frac{4.41+62.41+47.61+228.01+0.01+26.01+1.21+285.61+4.41+37.21}{10}$$

$$= 69.69.$$

So $\frac{a+b}{2} = 37.1 \qquad \frac{(b-a)^2}{12} = 69.69$

$a+b = 74.2 \qquad b-a = 28.9185$

$a = 22.64075 \qquad b = 51.55925$

Question 2 – 10 points

Consider the following game of chance where based on the outcome of a coin toss (of a biased coin) you win either \$5 (on getting heads) or \$10 (on getting tails). The outcomes for 8 games are as follows: \$5, \$10, \$5, \$10, \$10, \$5, \$10, \$10. From this find the probability of getting heads.

If X represents one game winnings

$$p(x) = \begin{cases} p, & \cancel{X} = 5 \\ 1-p, & X = 10 \end{cases}$$

$$P(X) = P(X_1 = 5, X_2 = 10, X_3 = 5 \ldots X_8 = 10) = (p)(1-p)(p)\cdot(1-p)(1-p)\cdot p\cdot(1-p)(1-p)$$

$$= p^3 (1-p)^5$$

To find p, $\frac{\partial}{\partial p} \ln p^3 (1-p)^5 = 0$

$$\frac{\partial}{\partial p}\left[\ln p^3 + \ln (1-p)^5\right] = 0$$

$$\frac{\partial}{\partial p}\left[3 \ln p + 5 \ln(1-p)\right] = 0$$

$$\frac{3}{p} + \frac{5}{1-p}(-1) = 0$$

$$\frac{3}{p} = \frac{5}{1-p}$$

$3 - 3p = 5p$ $\quad p = \frac{3}{8} = 0.375$

Name: ______________________________

Question 3 – 20 points

Students are polled to evaluate their opinion of two new University policies (A and B). Of the 600 students polled, 240 were in support of policy A. Of the 800 students polled, 320 were in support of policy B. Give the 90% confidence interval for the percentage of students in support of policy A and in support of policy B. Also give the 90% confidence interval for the difference in support for both these policies.

Is there enough statistically significant evidence to support that both the proposals are equally supported by the student population.

Since $n_A = 600$, $n_B = 800$ we can use Z statistic

$\alpha = 0.1.$

Confidence interval for $P_A = \dfrac{240}{600} \pm Z_{0.05} \sqrt{\dfrac{\frac{240}{600} \times \frac{360}{600}}{600}}$

$= 0.4 \pm 1.645 \sqrt{\dfrac{0.4 \times 0.6}{600}}$

$= 0.4 \pm 0.0329$

Confidence interval for $P_B = \dfrac{320}{800} \pm Z_{0.05} \sqrt{\dfrac{\frac{320}{800} \times \frac{480}{800}}{800}}$

$= 0.4 \pm 0.02849$

Confidence interval for $P_A - P_B$

$$= 0.4 - 0.4 \pm 1.645 \sqrt{\frac{0.4 \times 0.6}{600} + \frac{0.4 \times 0.6}{800}}$$

$$= 0 \pm 0.0435$$

Pooled proportion

$$\hat{p} = \frac{600 \times 0.4 + 800\ 0.4}{600 + 800}$$

$$= 0.4.$$

$H_A : P_A \neq P_B$

$H_0 : P_A = P_B$

$$Z_{obs} = \frac{0.4 - 0.4.}{\sqrt{0.4(1-0.4)\left(\frac{1}{600} + \frac{1}{800}\right)}} = 0$$

$$|Z_{obs}| = 0 < Z_{0.05}$$

So accept H_0.

Name: ______________________________

Question 4 – 20 points

You are interested in the difference of the starting salaries of PhD and MS graduates. You obtain the starting salaries of 3 PhD graduates (in thousands of $): 120, 85, 90. You also obtain the starting salaries of 4 MS graduates (in thousands of $): 65, 70, 80, 75. Give the 95% confidence interval for difference in average salaries. With 2% significance, can you say that a PhD graduate's average salary is $25,000 more than that of a MS graduate.

Since number of samples is small, we use t-distribution.

Let μ_x : Average PhD Salary

μ_y : Average MS Salary

$H_0 : \mu_x - \mu_y = 25 \qquad H_A : \mu_x - \mu_y \neq 25$

Test statistic $t = \dfrac{\bar{X} - \bar{y} - 25}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}}$

Degrees of Freedom

$$V = \frac{\left(\frac{s_x^2}{n} + \frac{s_y^2}{m}\right)}{\frac{s_x^4}{n^2(n-1)} + \frac{s_y^4}{m^2(m-1)}}$$

$\bar{X} = 98.3 \qquad \bar{y} = 72.5 \qquad s_x = 18.92974_9$

$s_y = 6.45497$

$$\nu = \frac{(129.861666)^2}{7133.55626 + 36.16893}$$

$$= \cancel{0.018775} \; 2.352119 \approx 2.$$

rounding down, $\nu = 2$

$$t = \frac{98.3 - 72.5 - 25}{\sqrt{129.861666}}$$

$$= 0.070202$$

$$t_{\alpha/2} = t_{0.01} = \cancel{31.82} \; 6.965$$

Since $t < t_{\alpha/2}$

Accept H_0

So ~~there~~ can say at 2% significance there is enough evidence that difference in salary is \$25,000

Name: ______________________

Question 5 – 15 points

You want to know if the probability of a certain experiment succeeding p is either 0.7 or 0.5. To find that out you keep repeating the experiment multiple times each day till it succeeds once. The number of times you had to repeat the experiment each day before you got your first success is given by: 2, 7, 5, 4, 7. Based on the given data, which of these is more likely to be the correct value for p.

Prior probability

$$\pi(\theta) = \begin{cases} 0.5 & p \text{ is } 0.7 \\ 0.5 & p \text{ is } 0.5 \end{cases}$$

$f(x / p \text{ is } 0.7)$ is Geometric (0.7)

$f(x / p \text{ is } 0.5)$ is Geometric (0.5)

$$f\left(x / p \text{ is } 0.7\right) = (1-0.7)^1 (0.7)\ (1-0.7)^6 (0.7)\ (1-0.7)^4 (0.7)\ (1-0.7)^3 (0.7)\ (1-0.7)^6 (0.7)$$

$$= (0.3)^{20} (0.7)^5$$

$$f\left(x / p \text{ is } 0.5\right) = (0.5)^{20} (0.5)^5 = (0.5)^{25}$$

Marginal $m(x) = \pi(p \text{ is } 0.7)\, f\left(x / p \text{ is } 0.7\right) + \pi(p \text{ is } 0.5)\, f\left(x / p \text{ is } 0.5\right)$

$$(0.5)(0.3)^{20}(0.7)^{5}+(0.5)(0.5)^{25}$$

Posterior,

$$\pi\left(\text{p is 0.7}/x\right)=\frac{f\left(x/\text{p is 0.7}\right)\,\pi\left(\text{p is 0.7}\right)}{m(x)}$$

$$=\frac{(0.3)^{20}(0.7)^{5}\,(\cancel{0.5})}{(\cancel{0.5})(0.3)^{20}(0.7)^{5}+(0.5)^{25}(\cancel{0.5})}$$

$$=0.00019659$$

$$\pi\left(\text{p is 0.5}/x\right)=\frac{f\left(x/\text{p is 0.5}\right)\,\pi\left(\text{p is 0.5}\right)}{m(x)}$$

$$=\frac{(0.5)^{25}\,(\cancel{0.5})}{(\cancel{0.5})(0.3)^{20}(0.7)^{5}+(\cancel{0.5})(0.5)^{25}}$$

$$=0.99980341$$

p is most likely $\underline{\underline{0.5}}$

Name: ___________________________

Question 6 – 25 points

You think that the rent for a one-bedroom apartment in Arlington, Texas follows a normal distribution and are 95% sure (before collecting any data) that it is between [$650, $1500]. You survey 500 people who live in such apartments about their rent. Your results follow a normal distribution with mean $900 and standard deviation 300. What does your posterior distribution look like. Calculate the 95% HPD Credible Set for the rent for a one-bedroom apartment in Arlington. Also calculate a Bayesian Estimate and give me the posterior risk for the rent. Is it likely that the rent is more than $1200.

Assuming [650, 1500] is the 95% Credible set for the prior Normal (μ, τ)

$$\left[\mu - Z_{0.05/2}\,\tau ,\ \mu + Z_{0.05/2}\,\tau\right] = [650, 1500]$$

$$\mu - 1.96\,\tau = 650$$

$$\mu + 1.96\,\tau = 1500$$

$$\mu = 1075 \qquad \tau = 216.83673$$

$$\pi(\theta) = \text{Normal}(1075, 216.83673)$$

$$f(x/\theta) = \text{Normal}(900, 300) \qquad n = 500$$

Normal is conjugate to Normal

Posterior

$$\pi\left(\theta/x\right) = \left(\frac{\frac{500 \times 900}{300^2} + \frac{1075}{216.83673^2}}{500/300^2 + 1/216.83673^2}, \frac{1}{\sqrt{500/300^2 + 1/216.83673^2}} \right)$$

$$= \left(\frac{5.022863}{0.0055768}, \frac{1}{\sqrt{0.0055768}} \right)$$

$$= \left(900.67126, 13.39082 \right)$$

95% HPD Credible set is.

$$900.67126 \pm 1.96 \times 13.39082$$

$$\left[874.42525, 926.91726 \right]$$

 or $[\$874.42, \$926.91]$

Name: ____________________________

Bayesian Estimate of average rent

$\bar{\Theta}_B = \$900.67$

$H_0 : \theta > 1200 \qquad H_A : \theta \leq 1200$

$\pi(H_A/x) = \pi(\theta \leq 1200/x) = P\left(Z \leq \frac{1200 - 900.67126}{13.39082}\right)$

$= P(Z \leq 22.35327)$

$= 1$

$\pi(H_0/x) = 1 - \pi(H_A/x) = 0$

Reject H_0 & accept H_A

Average rent is less than \$1200

Posterior risk $= 0 \cdot \omega_1 + 1 \cdot \omega_2 = \omega_2$

when ω_2 is loss of choosing H_A incorrectly

Question 7 (OPTIONAL) – 10 points (Extra Credit)

It is imperative that resistance of a critical electronic component not deviate from the specs. The manufacturer claims that the standard deviation of their components will not exceed 0.05 ohms. You test 25 samples and find the sample standard deviation to be 0.07 ohms. At 1% significance level, can you accept the manufacturers claim?

$H_0 : \sigma^2 = (0.05)^2 \qquad H_A : \sigma^2 > (0.05)^2$

~~test statistic~~

$$\chi^2_{obs} = \frac{24 \times (0.07)^2}{(0.05)^2} = 47.04$$

$$\chi^2_{\alpha} = \chi^2_{0.01} \quad \text{with } \nu = 24.$$

$$= 43.$$

$$\chi^2_{obs} > \chi^2_{0.01}$$

So reject H_0 & accept H_A

reject the manufacturers claim.

Name: ___________________________

SCRATCH