Assignment 3
MATH414
Statistics

INSTRUCTIONS

- Write your assignments with a pen on paper and submit it in the physical form.
- The use of ChatGPT or any other generative AI or, in fact, any "calculator" such as "WolframAlpha" is strictly forbidden. If you use any of these, you will get zero in the entire Assignment. The exception to this rule is when you want to check a calculation once you have performed it.
- You are welcome to look up any theory or definitions from any source you like but you must cite it.
- Write your answers cleanly.
- You need to fully write the R commands that you use.

QUESTIONS

(1) (Athreya, Sarkar & Tanner, pp331) Let $\lambda > 0$ and $X_1, X_2, \ldots, X_n$ be an i.i.d. sample from a population with Exponential(λ) distribution.
   (a) Find the likelihood function.
   (b) Find the maximum likelihood estimate for λ. Use the second derivative test to verify that this is indeed a maximum.

(2) (Athreya, Sarkar & Tanner, pp331) Let $X_1, X_2, \ldots, X_n$ be an i.i.d. sample from a population with Poisson(λ) distribution, where $\lambda > 0$.
   (a) Find the likelihood function $L(\lambda; X_1, X_2, \ldots, X_n)$.
   (b) Prove that if at least one of X_j values is non-zero, then the maximum likelihood estimate for λ is $\bar{X}$.
   (c) Prove that if all of the X_j values are zero, then $L(\lambda; X_1, X_2, \ldots, X_n)$ has no maximum value for $\lambda > 0$.

(3) (Athreya, Sarkar & Tanner, pp331) Let $0 < p < 1$ and $X_1, X_2, \ldots, X_n$ be an i.i.d. sample from a population with Geometric(p) distribution.
   (a) Find the likelihood function $L(p; X_1, X_2, \ldots, X_n)$.
   (b) Prove that the maximum likelihood estimate for p is $1/\bar{X}$.

(4) (Hogg, McKean, Craig, pp244) Let the observed value of the mean $\bar{X}$ and of the sample variance of a random sample of size 20 from a distribution that is $N(\mu, \sigma^2)$ be 81.2 and

26.5 respectively. Find $90\%, 95\%$ and 99% confidence intervals for μ. How does the length of the confidence interval increase as the confidence increases.

(5) (Hogg, McKean, Craig, pp244) Suppose we assume that $X_1, X_2, \ldots, X_n$ is a random sample from a $\Gamma(1, \theta)$ distribution.

(a) Show that the random variable $2/\theta \sum_{i=1}^{n} X_i$ has a χ^2-distribution with $2n$ degrees of freedom.

(b) Using the random variable in part(a) as a pivot random variable, find a $(1-\alpha)100\%$ confidence interval for θ.

(6) (Hogg, McKean, Craig, pp245) Let $\bar{X}$ be the mean of a random sample of size n from a distribution that is $N(\mu, 9)$. Find n such that $P(\bar{X} - 1 < \mu < \bar{X} + 1) = 0.90$.

(7) (Athreya, Sarkar & Tanner, pp335) Let $\lambda > 0$ and $X_1, X_2, \ldots, X_n$ be an i.i.d. sample from an Exponential(λ) distribution with rate λ.

(a) Let $T(X_1, X_2, \ldots, X_n, \lambda) = n\lambda\bar{X}$. Prove that T follows a Gamma$(n, 1)$ distribution.

(b) Suppose $n = 15$ and $\bar{X} = 2$. Find a 95% confidence interval.

(8) (Hogg, McKean, Craig, pp247) Let two independent random variables, each of size 10 from two normal distributions $N(\mu_1, \sigma^2)$ and $N(\mu_2, \sigma^2)$ yield $\bar{x} = 4.8$, $s_x^2 = 8.64$, $\bar{y} = 5.6$, $s_y^2 = 7.88$. Find a 95% confidence interval for $\mu_x - \mu_y$.

(9) (Hogg, McKean, Craig, pp247-8) In the previous problem, when the variances are distinct and unknown, it is difficult to find a test statistic. However, if the ratio of the two variances is known, then a test statistic can be found. Let $X_1, X_2, \ldots, X_9$ and $Y_1, Y_2, \ldots, Y_n$ be two independent random samples the respective normal distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$. It is given that $\sigma_1^2 = 3\sigma_2^2$, but σ_2^2 is unknown. Define a pivot variable that has a t-distribution that can be used to find a 95% confidence interval for $\mu_1 - \mu_2$.

(10) (Pruim pp121) A child's game includes a spinner with four colours on it. Each color is one quarter of the circle. You want to test the spinner to see if it is fair, so you decide to spin the spinner 50 times and count the number of blues. You do this and record 8 blues.

Carry out a hypothesis test, carefully showing the four steps. Do it "by hand" (using R but not `binom.test()`). Then check your work using `binom.test()`.

(11) (Pruim pp121) Fred wants to know whether his cat, Gus, has a preference for one paw or uses both paws equally. He dangles a ribbon in front of the cat and records which paw Gus uses to bat at it. He does this 10 times, and Gus bats at the ribbon with his right paw 8 times and his left paw 2 times. Then Gus gets bored with the experiment and leaves. Can Fred conclude that Gus is right-pawed, or could this result have occurred simply due to chance under the null hypothesis that Gus bats equally often with each paw?

(12) (Pruim pp121) *Roptrocerus xylophagorum* is a parasitoid of bark beetles. To determine what cues these wasps use to find their hosts, researchers placed female wasps in the base of a Y-shaped tube, with a different odor in each arm of the Y, and recorded the number of wasps that entered each arm.

(a) In one experiment, one arm of the Y had the odor of bark being eaten by adult beetles, while the other arm of the Y had the odor of bark being eaten by larval beetles. Ten wasps entered the area with the adult beetles, while 17 entered the area with the larval beetles. Calculate a p-value and draw a conclusion.

(b) In another experiment that compared infested bark with a mixture of infested and uninfested bark, 36 wasps moved towards the infested bark, while only 7 moved towards the mixture. Calculate a p-value and draw a conclusion.

(13) (Pruim pp122) Suppose you are going to test a coin to see if it is fair.

(a) You decide to flip it 200 times and to conduct a binomial test with the data you collect. Suppose the coin is actually biased so that it comes up heads 55% of the time. What is the probability that your test will have a p-value less than 0.05?

(b) How does your answer change if you flip the coin 400 times instead?

(c) How many flips must you perform if you want a 90% chance of detecting a coin that comes up heads 55% of the time? (Assume here again that we will reject the null hypothesis if the p-value is less than 0.05.)

(d) Express the results in the previous exercise using the terminology of statistical power.

[Hint: How many heads would it take for you to get a p-value less than 0.05?]