King Saud University

College of Science

Statistics and Operations Research Department

Stat 319

Theory of Statistics (2)

Exercises

References:

1. Introduction to Mathematical Statistics, Sixth Edition, by R. Hogg, J. McKean, and A. Craig, Prentice Hall.
2. Introduction to Theory of Statistics, A. Mood, F. Graybill and B. Boes, McGrow-Hill.
3. Introduction to Mathematical Statistics, Fourth Edition, by Hogg and Craig, Macmillan Publishing Co., Inc.
4. Statistical Inference, second Edition, by G. Casella and R. Berger, Duxbury.

By

Dr. Samah Alghamdi

Confidence Intervals

1. Let $\bar{X}$ be the mean of a random sample from the exponential distribution, $Exp(\theta)$.
   (a) Show that $\bar{X}$ is an unbiased point estimator of θ.
   (b) Using the MGF technique determine the distribution of $\bar{X}$.
   (c) Use (b) to show that $Y = 2n\bar{X}/\theta$ has a χ^2 distribution with 2n degrees of freedom.
   (d) Based on Part (c), find a 95% confidence interval for θ if $n = 10$. Hint: Find c and d such that $P\left(c < \frac{2n\bar{X}}{\theta} < d\right) = 0.95$ and solve the inequalities for θ.
2. Let $X_1, \dots, X_n$ be a random sample from the $\Gamma(2, \theta)$ distribution, where θ is unknown. Let $Y = \sum_{i=1}^{n} X_i$.
   (a) Find the distribution of Y and determine c so that cY is an unbiased estimator of θ.
   (b) If $n = 5$, show that $P\left(9.59 < \frac{2Y}{\theta} < 34.2\right) = 0.95$.
   (c) Using Part (b), show that if y is the value of Y once the sample is drawn, then the interval $\left(\frac{2y}{34.2}, \frac{2y}{9.59}\right)$ is a 95% confidence interval for θ.
   (d) Suppose the sample results in the values,
   44.8079 1.5215 12.1929 12.5734 43.2305
   Based on these data, obtain the point estimate of θ as determined in Part (a) and the computed 95% confidence interval in Part (c). What does the confidence interval mean?
3. Let the observed value of the mean $\bar{X}$ of a random sample of size 20 from a distribution that $N(\mu, 80)$ be 81.2. Find a 95% confidence interval for μ.
4. 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 $P(\bar{X} - 1 < \mu < \bar{X} + 1) = 0.90$, approximately.
5. Let a random sample of size 17 from the normal distribution $N(\mu, \sigma^2)$ yield $\bar{x} = 4.7$ and $s^2 = 5.76$. Determine a 90% confidence interval for μ.
6. Let $\bar{X}$ denote the mean of a random sample of size n from a distribution that has mean μ and variance $\sigma^2 = 10$. Find n so that the probability is approximately 0.954 that the random interval $\left(\bar{X} - \frac{1}{2} < \mu < \bar{X} + \frac{1}{2}\right)$ includes μ.
7. Let Y be $Bin(300, p)$. If the observed value of Y is $y = 75$, find an approximate 90% confidence interval for p.
8. Let $\bar{x}$ be the observed mean of a random sample of size n from a distribution having mean μ and known variance σ^2. Find n so that $\bar{x} - \frac{\sigma}{4}$ to $\bar{x} + \frac{\sigma}{4}$ is an approximate 95% confidence interval for μ.
9. Let $X_1, \dots, X_n$ be a random sample from $N(\mu, \sigma^2)$, where both parameters μ and σ^2 are unknown. A confidence interval for σ^2 can be found as follows:
   We know that $(n-1)S^2/\sigma^2$ is a random variable with a $\chi^2_{(n-1)}$ distribution. Thus, we can find constants a and b so that $P((n-1)S^2/\sigma^2 < b) = 0.975$ and $P(a < (n-1)S^2/\sigma^2 < b) = 0.95$.
   (a) Show that this second probability statement can be written as
   $P((n-1)S^2/b < \sigma^2 < (n-1)S^2/a) = 0.95$.

(b) If $n = 9$ and $s^2 = 7.93$, find a 95% confidence interval for σ^2.

10. Let $X_1, \dots, X_n$ be a random sample from gamma distributions with known parameter $\alpha = 3$ and unknown $\beta > 0$. Discuss the construction of a confidence interval for β.
Hint: What is the distribution of $2\sum_{i=1}^{n}\frac{X_i}{\beta}$? Follow the procedure outlined in Exercise 9.
11. When 100 tacks were thrown on a table, 60 of them landed point up. Obtain a 95% confidence interval for the probability that a tack of this type will land point up. Assume independence.
12. Let two independent random samples, each of size 10, from two normal distributions $N(\mu_1, \sigma^2)$ and $N(\mu_2, \sigma^2)$ yield $\bar{x} = 4.8$, $s_1^2 = 8.64$, $\bar{y} = 5.6, s_2^2 = 7.88$. Find a 95% confidence interval for $\mu_1 - \mu_2$.
13. Let two independent random variables, Y_1 and Y_2, with binomial distribution that have parameters $n_1 = n_2 = 100, p_1$ and p_2, respectively, be observed to be equal to $y_1 = 50$ and $y_2 = 40$. Determine an approximate 90% confidence interval for $p_1 - p_2$.
14. Let $\bar{X}$ and $\bar{Y}$ be the means of two independent random samples, each of size n, from the respective distributions $N(\mu_1, \sigma^2)$ and $N(\mu_2, \sigma^2)$, where the common variance is known. Find n such that $P\left(\bar{X} - \bar{Y} - \frac{\sigma}{5} < \mu_1 - \mu_2 < \bar{X} - \bar{Y} + \frac{\sigma}{5}\right) = 0.90$.
15. If 8.6, 7.9, 8.3, 6.4, 8.4, 9.8, 7.2, 7.8, 7.5 are the observed values of a random sample of size 9 from a distribution that is $N(8, \sigma^2)$, construct a 90% confidence interval for σ^2.
16. Let $X_1, \dots, X_n$ and $Y_1, \dots, Y_m$ be two independent random samples from the respective normal distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, where the four parameters are unknown. To construct a confidence interval for the ratio, $\frac{\sigma_1^2}{\sigma_2^2}$, of the variances, from the quotient of the two independent chi-square variables, each divided by its degrees of freedom, namely $F = \frac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2}$, where S_1^2 and S_2^2 are the respective sample variances.
    (a) What kind of distribution does F have?
    (b) From the appropriate table, a and b can be found so that $P(F < b) = 0.975$ and $P(a < F < b) = 0.95$.
    (c) Rewrite the second probability statement as $P\left(a\frac{S_1^2}{S_2^2} < \frac{\sigma_1^2}{\sigma_2^2} < b\frac{S_1^2}{S_2^2}\right) = 0.95$. The observed values, s_1^2 and s_2^2, can be inserted in these inequalities to provide a 95% confidence interval for $\frac{\sigma_1^2}{\sigma_2^2}$.
17. Let two independent random samples of sizes $n = 16$, $m = 10$, taken from two independent normal distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, respectively, yield $\bar{x} = 3.6$, $s_1^2 = 4.14$, $\bar{y} = 13.6, s_2^2 = 7.26$. Fin a 90% confidence interval for $\frac{\sigma_2^2}{\sigma_1^2}$ when μ_1 and μ_2 are unknown.
18. Find a 90 percent confidence interval for the mean of a normal distribution with $\sigma = 3$ given the sample (3.3, -0.3, -0.6, -0.9). What would be the confidence interval if σ were unknown?
19. The breaking strengths in pounds of five specimens of manila rope of diameter $\frac{3}{16}$ inch were found to be 660, 460, 540, 580 and 550.

(a) Estimate the mean breaking strength by a 95% confidence interval assuming normality.
(b) Estimate σ^2 by a 90% confidence interval; also σ.

20. A sample was drawn from each of five populations assumed to be normal with the same variance. The values of $(n-1)S^2 = \sum(X_i - \bar{X})$ and n, the sample size, were

S^2:	40	30	20	42	50
n:	6	4	3	7	8

Find 98% confidence interval for the common variance.

21. To test two promising new lines of hybrid corn under normal farming conditions, a seed company selected eight farms at random in Iowa and planted both lines in experimental plots on each farm. The yields (converted to bushels per acre) for the eight locations were

Line A:	86	87	56	93	84	93	75	79
Line B:	80	79	58	91	77	82	74	66

Assuming that the two yields are jointly normally distributed, estimate the difference between the mean yields by a 95% confidence interval.

22. Let $\bar{X}$ denote the mean of a random sample of size 25 from a gamma distribution with $\alpha = 4$ and $\beta > 0$. Use the Central Limit Theorem to find an approximate 0.954 confidence interval for μ, the mean of the gamma distribution. Hint: Use the random variable
$\frac{(\bar{X}-4\beta)}{\sqrt{4\beta/25}} = \frac{5\bar{X}}{2\beta} - 10.$

23. Let S_1^2 and S_2^2 denote, respectively, the variances of random samples, of sizes n and m, from two independent distributions that are $N(\mu_1, \sigma^2)$ and $N(\mu_2, \sigma^2)$.
(a) Derive $(1-\alpha)100\%$ confidence interval for the common unknown variance σ^2.
(b) Find 99% confidence interval for σ^2 if $n = 12, s_1^2 = 4.74, m = 15, s_2^2 = 5.66$.

Type I Error, Type II Error and Power Function

1. Let X have a binomial distribution with the number of trials $n = 10$ and with p either 1/4 or 1/2. The simple hypothesis $H_0: p = \frac{1}{2}$ is rejected, and the alternative simple hypothesis $H_1: p = \frac{1}{4}$ is accepted, if the observed value of X_1, a random sample of size 1, is less than or equal to 3. Find the significance level and the power of the test.
2. Let $X_1, X_2 \dots., X_{25}$ be a random sample from $N(\mu, 16)$. If the test: Reject $H_0: \mu = 1$ and accept $H_1: \mu = 3$ when $\bar{x} \geq 2$. Find the Type I error and the power function of the test.
3. Let X_1, X_2 be a random sample of size $n = 2$ from the distribution having pdf form$f(x;\theta) = \frac{1}{\theta} e^{-x/\theta}, x > 0, \theta > 0$. We reject $H_0: \theta = 2\ vs\ H_1: \theta = 1$ if the observed values of X_1, X_2, say x_1, x_2 are such that
$$\frac{f(x_1, 2) f(x_2, 2)}{f(x_1, 1) f(x_2, 1)} \leq \frac{1}{2}$$
Here $\Omega = \{\theta: \theta = 1,2\}$. Find the significance level of the test.
4. Let X have a Poisson distribution with mean θ. Consider the simple hypothesis $H_0: \theta = \frac{1}{2}$ and the alternative composite hypothesis $H_1: \theta < \frac{1}{2}$. Thus $\Omega = \left\{\theta: 0 < \theta \leq \frac{1}{2}\right\}$. Let $X_1, \dots., X_{12}$ denote a random sample of size 12 from this distribution. We reject H_0 if and only if observed value of $Y = X_1 + \cdots. + X_{12} \leq 2$. If $\gamma(\theta)$ is the power function of the test, find the power $\gamma\left(\frac{1}{2}\right), \gamma\left(\frac{1}{3}\right), \gamma\left(\frac{1}{4}\right), \gamma\left(\frac{1}{6}\right)$ and $\gamma\left(\frac{1}{12}\right)$. What is the significance level of the test?
5. Let $X_1, X_2 \dots., X_8$ be a random sample of size $n = 8$ from a Poisson distribution with mean μ. Reject the simple null hypothesis $H_0: \mu = 0.5$ and accept $H_1: \mu > 0.5$ if the observed sum $\sum_{i=1}^{8} x_i \geq 8$.
   (a) Compute the significance level α of the test.
   (b) Find the power function $\gamma(\mu)$ of the test as a sum of Poisson probabilities.
   (c) Using the Poisson Table, determine $\gamma(0.75), \gamma(1)$ and $\gamma(1.25)$.
6. Let X be a Bernoulli random variable with probability of success p. suppose we want to test at size α, $H_0: p = 0.7\ vs\ H_1: p = 0.5$. Reject H_0 if $\sum_{i=1}^{20} x_i \leq c$. Find c if the size of test is equal to 0.048. Then, find the power function of the test.
7. Let $X_1, \dots, X_n$ is normally distributed with mean μ and variance 100. Reject $H_0: \mu = 75\ vs\ H_1: \mu = 77$ if $\sum_{i=1}^{n} x_i > nc$. Determine n and c so that the power function $\gamma(\mu)$ of the test has the values $\gamma(75) = 0.159$ and $\gamma(77) = 0.841$.
8. Let X have a pdf of the form $f(x, \theta) = \theta x^{\theta - 1}, 0 < x < 1$, zero elsewhere, where $\theta \in \{\theta: \theta = 1, 2\}$. To test the simple hypothesis $H_0: \theta = 1$ against the alternative simple hypothesis $H_1: \theta = 2$, use a random sample X_1, X_2 of size $n = 2$ and define the critical region to be $C = \left\{(x_1, x_2): \frac{3}{4} \leq x_1 x_2\right\}$. Find the power function of the test.

9. Let $Y_1 < Y_2 < Y_3 < Y_4$ be the order statistics of a random sample of size $n = 4$ from a distribution with pdf $f(x;\ \theta) = {}^{1}\!/_{\theta}, 0 < x < \theta$, zero elsewhere, where $0 < \theta$. The hypothesis $H_0: \theta = 1$ is rejected and $H_1: \theta > 1$ is accepted if the observed $Y_4 \geq c$.
   (a) Find the constant c so that the significance level is $\alpha = 0.05$.
   (b) Determine the power function of the test.
10. Let X be a single observation from the density $f(x, \theta) = (2\theta x + 1 - \theta), 0 < x < 1$, where $-1 \leq \theta \leq 1$. To test $H_0: \theta = 0$ vs $H_1: \theta > 0$, the following procedure was used: Reject H_0 if X exceeds $\frac{1}{2}$. Find the power and the size of the test
11. Let $X_1, \dots, X_n$ denote a random sample from $f(x;\ \theta) = \left({}^{1}\!/_{\theta}\right), 0 < x < \theta$, and let $Y_1, \dots, Y_n$ be the corresponding ordered sample. To test $H_0: \theta = \theta_0$ versus $H_1: \theta \neq \theta_0$, the following test was used: Accept H_0 if $\theta_0\left(\sqrt[n]{\alpha}\right) \leq Y_n \leq \theta_0$; otherwise reject. Find the power function of this test.
12. Let $X_1, \dots, X_n$ be a random sample of size n from $f(x;\ \theta) = \theta^2 x e^{-\theta x}\ x > 0$. In testing $H_0: \theta = 1$ versus $H_1: \theta > 1$ for $n = 1$ (a sample of size 1) the following test was used: Reject H_0 if and only if $X_1 \leq 1$. Find the power function and size of the test.

Testing Hypotheses

Case 1: Simple Hypotheses (Neyman-Pearson Lemma)

1. Let the random variable X have the pdf $f(x;\theta) = \left({}^{1}\!/_{\theta}\right)e^{-x/\theta}, 0 < x < \infty$, zero elsewhere. Consider the simple hypothesis $H_0: \theta = 2$ and the alternative hypothesis $H_1: \theta = 4$. Let $X_1,\ X_2$ denote a random sample of size 2 from this distribution. Show that the best of H_0 against H_1 may be carried out by use of the statistic $X_1 +\ X_2$.
2. Let $X_1, X_2, \ldots\ldots X_{10}$ be a random sample of size 10 from a normal distribution $N(0, \sigma^2)$. Find a best critical region of size $\alpha = 0.05$ for testing $H_0: \sigma^2 = 1$ against $H_1: \sigma^2 = 2$.
3. If $X_1, X_2, \ldots\ldots X_n$ is a random sample from a distribution having pdf of the form $f(x;\theta) = \theta x^{\theta-1}, 0 < x < 1$, zero elsewhere, show that a best critical region for testing $H_0: \theta = 1$ against $H_1: \theta = 2$ is $C = \{(x_1, x_2, \ldots\ldots x_n): c \leq \prod_{i=1}^{n} x_i\}$.
4. Let $X_1, X_2, \ldots\ldots X_{10}$ be a random sample from a distribution that is $N(\theta_1, \theta_2)$. Find a best test of the simple hypothesis $H_0: \theta_1 = 0, \theta_2 = 1$ against the alternative simple hypothesis $H_1: \theta_1 = 1,\ \theta_2 = 4$.
5. Let $X_1, X_2, \ldots\ldots X_n$ denote a random sample from a normal distribution $N(\theta, 100)$. Show that $C = \{(x_1, x_2, \ldots\ldots x_n): c \leq\ \bar{x}\}$ is a best critical region for testing $H_0: \theta = 75$ against $H_1: \theta = 78$. Find n and c so that

$$P[(X_1, X_2, \ldots., X_n) \in C;\ H_0] = P(\bar{X} \geq c;\ H_0) = 0.05 \text{ and}$$

$$P[(X_1, X_2, \ldots., X_n) \in C;\ H_1] = P(\bar{X} \geq c;\ H_1) = 0.90, \text{ approximately.}$$

6. If $X_1, X_2, \ldots\ldots X_n$ is a random sample from a beta distribution with parameters $\alpha = \beta = \theta > 0$, find a best critical region for testing $H_0: \theta = 1$ against $H_1: \theta = 2$.
7. Let $X_1, X_2, \ldots\ldots X_n$ be iid with pmf $f(x;p) = p^x(1-p)^{1-x}, x = 0, 1$, zero elsewhere. Show that $C = \{(x_1, x_2, \ldots\ldots x_n): \sum_1^n x_i \leq c\}$ is a best critical region for testing $H_0: p = \frac{1}{2}$ against $H_1: p = \frac{1}{3}$. Use the Central Limit Theorem to find n and c so that approximately $P(\sum_1^n X_i \leq c;\ H_0) = 0.10$ and $P(\sum_1^n X_i \leq c;\ H_1) = 0.80$.
8. Let $X_1, X_2, \ldots\ldots X_{10}$ denote a random sample of size 10 from a Poisson distribution with mean θ. Show that the critical region C defined by $\sum_1^{10} x_i \geq c$ is a best critical region for testing $H_0: \theta = 0.1$ against $H_1: \theta = 0.5$. Determine, for this test, c and the power at $\theta = 0.5$ when the significance level $\alpha = 0.08$.
9. Let X be a random variable has the pdf

$$f(x) = \frac{\beta}{x^{\beta+1}}\ , \qquad x \geq 1, \qquad \beta > 0$$

Find the best critical region of size α for testing

$$H_0: \beta = \beta_0 \ \text{ versus } \ H_1: \beta = \beta_1\,,\ \beta_1 > \beta_0.$$

10. Let $X_1, X_2, \dots\dots X_n$ be independent random variables have $N(\mu, \sigma^2)$. Find the best critical region of size α for testing:
(i) H_0: $\mu = \mu_0$ against H_1: $\mu = \mu_1$ when $\mu_0 > \mu_1$ and σ^2 is known.
(ii) H_0: $\sigma^2 = \sigma_0^2$ against H_1: $\sigma^2 = \sigma_1^2$ when $\sigma_0^2 < \sigma_1^2$ and μ is known.

Case 2: Simple Hypothesis against One-Sided Hypothesis

1. Consider a normal distribution of the form $N(\theta, 4)$. Test the simple hypothesis $H_0: \theta = 0$ against the alternative composite hypothesis $H_1: \theta > 0$ and find the best rejection region.
2. Let $X_1, X_2, \dots, X_n$ be a random sample from a distribution with pdf $f(x; \theta) = \theta x^{\theta - 1}, 0 < x < 1$, zero elsewhere, where $\theta > 0$. Calculate the best critical region to reject $H_0: \theta = 6$ versus $H_1: \theta < 6$.
3. Let X have the pdf $f(x; \theta) = \theta^x (1 - \theta)^{1-x}, x = 0, 1$, zero elsewhere. We test $H_0: \theta = \frac{1}{2}$ against $H_1: \theta < \frac{1}{2}$ by taking a random sample $X_1, X_2, \dots, X_5$ of size $n = 5$. Find the best critical region of size α.
4. Suppose that $X_1, X_2, \dots, X_n$ is a random sample of size *n* from exponential population with parameter $\frac{1}{\theta}$. Determine the best rejection region of size α for $H_0: \theta = \theta_0$ against $H_1: \theta > \theta_0$.
5. If $X \sim Gamma\left(2, \frac{1}{\theta}\right)$. Find the best critical region for $H_0: \theta = 1$ against $H_1: \theta < 1$.
6. Let X be a random sample whose probability mass function is binomial distribution with parameters $n = 10$ and *p*. Test the hypotheses $H_0: p = 0.25$ against $H_1: p > 0.25$ and find the best critical region of size α of this test.