99
(9)

Department of Mathematics & Statistics		
Course	**Number**	**Section(s)**
Statistics	250	AA
Examination	**Date**	**Pages**
Final	December 2013	2
Instructor		
L. Popovic		

Special Instructions:

- ▷ Answer all questions and show your work in clear steps leading to the final answer.
- ▷ No aids are allowed other than the provided sheets and an approved calculator.

---

1. Let Y_1, Y_2 be random variables with joint density function

$$f(y_1, y_2) = \begin{cases} \frac{1}{8} y_1 e^{-(y_1+y_2)/2} & \text{for } y_1 > 0, y_2 > 0 \\ 0 & \text{otherwise} \end{cases}$$

   a. (4 pts) Calculate $P(Y_1 > 1, Y_2 > 1)$.

   b. (4 pts) Find the marginal densities of Y_1 and of Y_2.

   b. (4 pts) Find the conditional density of $Y_1|Y_2 = y_2$.

   c. (2 pts) Are Y_1 and Y_2 independent? (*Justify your answer with an explanation*).

---

2. Let Y_1, Y_2, Y_3 be independent random variables with the same density $f(y) = e^{-y}$, $y > 0$.

   a. (6 pts) Find the density of $U = \min(Y_1, Y_2, Y_3)$.

   b. (4 pts) What variable is U? Use this to find $E(U)$ and $V(U)$. (*Do not use integration*)

---

3. Let $Y_1, \ldots, Y_n$ be a sample from the same density $f(y) = \lambda e^{-\lambda y}, y \geq 0$ representing the weights of items. Let $U = \sum_{i=1}^{n} Y_i$ be the sum total of the weights in the sample.

   a. (4 pts) Find the density of U.

   b. (2 pts) Find its mean $E(U)$ and variance $V(U)$. (*Do not use integration of the density of U*)

   c. (4 pts) If $n = 100$ and $\lambda = 20$ estimate the probability $P(U > 5.5)$. (*Do not use integration*)

   b. (2 pts) What approximation have you used in order to calculate the probability in part c.?

---

4. Two groups of patients of 30 individuals each were tested in a medical study designed to test the effectiveness of a new drug. The first group of patients received a placebo, while the second group of patients received the new drug, and their blood pressure was measured a few hours later. The sample mean and sample standard deviation of first group's pressure were $\overline{y}_1 = 167.1, s_1^2 = 24.3$, and for the second group they were $\overline{y}_2 = 140.9, s_2^2 = 17.6$.

a. (6 pts) Construct a 95% confidence interval for the true difference between the mean blood pressure after taking the placebo and after taking the new drug.

b. (2 pts) What statistic did you use to construct the CI and why?

b. (2 pts) Would you say it is likely that the new drug lowers the blood pressure by at least 15 points?

---

5. Let $Y_1, \ldots, Y_n$ be a sample from density $f(y) = \lambda e^{-\lambda y}, y \geq 0$.

a. (6 pts) Find the Maximum Likelihood Estimator (MLE) for λ.

b. (6 pts) Find a sufficient statistic for λ.

c. (6 pts) Find a Moment of Methods estimator for λ.

---

6. Let $Y_1, \ldots, Y_n$ be a sample from density $f(y) = \frac{1}{\beta} e^{-\frac{y}{\beta}}, y \geq 0$. Consider the estimator $\widehat{\beta} = n \min(Y_1, \ldots, Y_n)$.

a. (8 pts) Is $\widehat{\beta}$ an unbiased estimator of β? (*Justify your answer with a calculation*).

b. (4 pts) Is $\widehat{\beta}$ a consistent estimator of β? (*Justify your answer with a calculation*).

d. (4 pts) Is $\widehat{\beta}$ more efficient than the estimator $\overline{Y}$? (*Justify your answer with a comparison*).

---

7. Medical researchers were concerned that the new drug causes more variation than usually observed in patients' blood pressures. They ran a second study with in which they particularly focused on the sample variances of the two groups of measurements. In the second study they used a group of 14 patients taking the placebo and observed sample variance $s_1 = 12.7$, and a group of 10 patients taking the new drug and observed a sample variance of $s_2 = 26.4$

a. (2 pts) What is the null hypothesis, and what is the appropriate alternative hypothesis?

b. (2 pts) What statistic do you need to use for the hypothesis test, and what distribution does it have?

a. (6 pts) If we are allowing a type I error of 0.05, would you reject the null hypothesis based on the above data?

---

8. We have the following data on levels of safety Y (in percentages) as a function of pollution levels X:

$$(x, y) = \{(0.1, 100),\ (0.2, 95),\ (0.3, 85),\ (0.4, 65),\ (0.5, 55)\}$$

a. (6 pts) Give the equation for the line that best describes the dependence of Y on X.

b. (2 pts) Give an estimate of the level of safety when the pollution levels are at a record high of 0.9?

c. (2 pts) Do you think the equation should be used for the pollution level of 0.9?
