Statistics 426: Final Exam

Moulinath Banerjee

April 28, 2004

Announcement: The final exam carries 80 points. Half of what you score contributes to your grade in the course.

(1) . (a) Let Z_1, Z_2, Z_3 be i.i.d. N(0,1) random variables. Let $R = \sqrt{Z_1^2 + Z_2^2 + Z_3^2}$. Find the density function of R. (Hint: Can you write down the density function of R^2?)

(b) Suppose that $X \sim \Gamma(\alpha, \lambda)$ and $Y \sim \Gamma(\beta, \lambda)$ where $\alpha, \beta, \lambda > 0$. Let

$$U = X + Y \text{ and } V = \frac{X}{X+Y}.$$

(i) Show that the joint density of (X, Y) is

$$f(x,y) = \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)}\, e^{-\lambda\,(x+y)}\, x^{\alpha-1}\, y^{\beta-1}\,, \quad x > 0, y > 0\,.$$

(ii) Compute the joint density of U and V. Deduce that they are independent and write down their marginal densities.

(c) Let X be a random variable with distribution function $F(x)$. Let $f(x)$ be the density function of X. Evaluate $\int_{-\infty}^{\infty} F(x)\, f(x)\, dx$. (Hint: How is $F(X)$ distributed?) (5 + 10 + 5 = 20 points)

(2) . (a) Consider two groups of patients, say Group A and Group B. The number of patients in each of these groups is 50. Group A patients were on a blood pressure drug for 4 weeks while Group B patients were on placebo. For each patient the systolic blood pressure has been recorded. We denote the measurements for Group A patients by $X_1, X_2, \ldots, X_{50}$ and the measurements for Group B patients by $Y_1, Y_2, \ldots, Y_{50}$.

Under appropriate conditions, it is reasonable to think of the X_i's as a random sample from a $N(\mu_1, \sigma^2)$ and the Y_i's as a random sample from a $N(\mu_2, \sigma^2)$ population. Also, the two populations may be assumed to be independent.

Describe how you construct a confidence interval for the difference in the mean blood pressure for these two groups, based on the above data.

Now suppose that you have been provided with the following statistics:

$$\overline{X} = 140,\ \overline{Y} = 150,\ \sum_{i=1}^{50} (X_i - \overline{X})^2 = 5250 \ \text{ and } \ \sum_{i=1}^{50} (Y_i - \overline{Y})^2 = 4800\,.$$

What is the confidence set for the mean difference in blood pressure?

(b) Let X be a random variable following a $\Gamma(4, \lambda)$ distribution. Find a level $1 - \alpha$ confidence interval for λ.

(c) Let $X_1, X_2, \ldots, X_5$ be i.i.d. observations from the uniform distribution on $[a, b]$ where $a < b$ are unknown quantities. Consider the following postulated model for the data:

$$a = \theta, b = \theta + 1 \ \text{ for some } \ \theta\,.$$

Let the observed values of the X_i's be $(-3, -2.3, -3.2, -2.8, -1)$. Would you trust the postulated model based on the above data? (10 + 5 + 5 = 20 points)

(3) (a) Consider random variables $Y_1, Y_2, \ldots, Y_n$ where

$$Y_i = \beta\, W_i + \epsilon_i\,.$$

Here $W_1, W_2, \ldots, W_n$ are *fixed constants* and $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ are i.i.d.$N(0, \sigma^2)$ random variables. Assume that σ^2 is known. Thus, the only unknown parameter is β.

(i) Show that the joint density of $Y_1, Y_2, \ldots, Y_n$ is

$$p(\underline{Y}, \beta) = \left(\frac{1}{2\,\pi\,\sigma^2}\right)^{n/2} \exp\left[-\frac{\sum_{i=1}^{n} (Y_i - \beta\, W_i)^2}{2\,\sigma^2}\right].$$

(ii) Compute $\hat{\beta}$, the MLE of β. Show that it is normally distributed with mean equal to β. Also find its variance.

(iii) The score function and its derivative are defined in the usual way as,

$$\dot{l}(\underline{Y}, \beta) = \frac{\partial}{\partial\,\beta} \log\, p(\underline{Y}, \beta) \ \text{ and } \ \ddot{l}(\underline{Y}, \beta) = \frac{\partial^2}{\partial\,\beta^2} \log\, p(\underline{Y}, \beta).$$

Compute the information $I(\beta) = E\,(\dot{l}^2(\underline{Y}, \beta)) = -E\,(\ddot{l}(\underline{Y}, \beta))$.

(iv) How does the variance of $\hat{\beta}$ compare to the information bound for unbiased estimators of β in this model? Is $\hat{\beta}$ the best unbiased estimator of β?

(b) Decide whether the following statements are true or false.

(i) The variance of the best unbiased estimator of a parameter is always equal to the information bound.

(ii) The MLE is always unbiased for the parameter. ((3 + 5 + 4 + 3) + 5 = 20 points.)

(4) Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with common distribution Exp$(1/\beta)$ where $\beta > 0$ is the parameter of interest. Thus the common density function is:

$$f(x, \beta) = \frac{1}{\beta} \exp(-x/\beta)\,,\; x > 0\,.$$

(a) Find $\mu_1 = E(X_1)$ and hence find a MOM estimate of β. Call this $\hat{\beta}_{MOM}$.

(b) Find the MLE of β. How are $\hat{\beta}_{MOM}$ and the MLE related?

(c) Show that

$$\sqrt{n}\,(\hat{\beta}_{MOM} - \beta) \to_d N(0, \beta^2)\,.$$

(d) Find an appropriate function g such that

$$\sqrt{n}\,(g(\hat{\beta}_{MOM}) - g(\beta)) \to_d N(0, 1)\,.$$

Hence find an approximate level $1 - \alpha$ confidence interval for β. (4 + 4 + 5 + 7 = 20 points)