University of California, Los Angeles
Department of Statistics

Statistics 100B

Instructor: Nicolas Christou

Exam 1
26 April 2013

Name: SOLUTIONS

Problem 1 (25 points)
Answer the following questions:

a. Suppose $\bar{X}$ is unbiased estimator of μ_X. Is $\bar{X}^2$ an unbiased estimator of μ_X^2?

$$E\bar{X}^2 = VAR(\bar{X}) + (E\bar{X})^2$$
$$= \frac{\sigma^2}{n} + \mu_X^2 \qquad \text{NO!}$$

b. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with $X_i \sim N(0, \sigma)$. Find $E\frac{X_i^2}{\sigma^2}$.

$$\left(\frac{X_i - 0}{\sigma}\right)^2 \sim \chi^2_1 \quad \Rightarrow \quad E\frac{X_i^2}{\sigma^2} = 1$$

c. Refer to part (b). What is the distribution of $\frac{X_1}{\sqrt{\frac{\sum_{i=2}^{n} X_i^2}{n-1}}}$? Please show all your work.

$$\frac{\frac{X_1 - 0}{\sigma}}{\sqrt{\frac{\sum_{i=2}^{n}\left(\frac{X_i - 0}{\sigma}\right)^2}{n-1}}} \sim t_{n-1}$$

d. Suppose the logarithm of a stock price in six months will follow the normal distribution with mean \$50 and standard deviation \$20. Find the probability that the stock price in six months will exceed \$79. Note: It is given that $ln(X) \sim N(50, 20)$.

$$P(\ln X > \ln 79) = P\left(Z > \frac{\ln 79 - 50}{20}\right) = P(Z > -2.28)$$
$$= P(Z < 2.28) = \underline{0.9887}$$

Problem 2 **(25 points)**
Answer the following questions:

a. Let $X \sim \Gamma(\alpha, \beta)$. Show that $\frac{2}{\beta}X \sim \chi^2_{2\alpha}$.

$$f(x) = \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)\beta^\alpha}$$

$$F(y) = P(Y \le y) = P\left(\frac{2}{\beta}X \le y\right) = P\left(X \le \frac{\beta y}{2}\right) = F_X\left(\frac{\beta y}{2}\right)$$

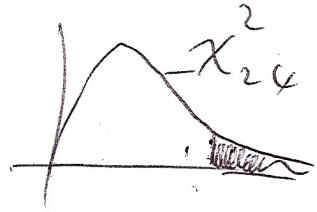
$$f_Y(y) = \frac{\beta}{2} f_X\left(\frac{\beta y}{2}\right) = \frac{\beta}{2}\,\frac{\left(\frac{\beta y}{2}\right)^{\alpha-1} e^{-\frac{\beta y}{2\beta}}}{\Gamma(\alpha)\beta^\alpha} = \frac{y^{\alpha-1} e^{-y/2}}{\Gamma(\alpha)\, 2^\alpha} \Rightarrow Y \sim \Gamma(\alpha, 2) \text{ OR } \chi^2_{2\alpha}$$

b. Refer to part (a). Suppose X_1, X_2, X_3, X_4 are i.i.d. random variables with $X_i \sim \Gamma(3, 4)$. Find c such that $P(\sum_{i=1}^4 X_i > c) = 0.05$.

$$P\left[\sum_{i=1}^4 X_i > c\right] = 0.05 \Rightarrow P\left[\frac{2}{4}\sum X_i > \frac{2}{4}c\right] = 0.05$$

$$\Rightarrow P\left[\frac{2}{4}X_1 + \dots + \frac{2}{4}X_4 > \frac{2}{4}c\right] = 0.05$$

$$\Rightarrow P\left[\chi^2_{24} > \frac{c}{2}\right] = 0.05$$



c. Let $X_1, X_2, \dots, X_n$ be i.i.d. random variables with $X_i \sim N(0, \sigma)$. Find $\hat{\sigma^2}$, the mle of σ^2.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{x^2}{\sigma^2}} \qquad L = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2}\sum \frac{x_i^2}{\sigma^2}}$$

$$\ln L = -\frac{n}{2}\ln 2\pi\sigma^2 - \frac{1}{2}\sum\frac{x_i^2}{\sigma^2}$$

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum x_i^2 = 0 \Rightarrow \hat{\sigma^2} = \frac{\sum x_i^2}{n}$$

d. Refer to part(c). Is $\hat{\sigma^2}$ MVUE of σ^2?

WORK WITH log-likelihood

$$\frac{\partial^2 \ln L}{\partial \sigma^{(2)}} = +\frac{n}{2\sigma^4} - \frac{1}{\sigma^6}\sum x_i^2$$

$$\frac{1}{-E\left(\frac{n}{2\sigma^4} - \frac{\sum x_i^2}{\sigma^6}\right)} = \frac{1}{-\left(\frac{n}{2\sigma^4} - \frac{n}{\sigma^4}\right)} = \frac{2\sigma^4}{n} \checkmark$$

YES MVUE.

$$VAR(\hat{\sigma^2}) = VAR\left(\frac{\sum x_i^2}{n}\right) = VAR\left[\frac{\sigma^2}{n}\sum\left(\frac{x_i}{\sigma}\right)^2\right] = \frac{2n\sigma^4}{n^2} = \frac{2\sigma^4}{n} \checkmark$$

Problem 3 (25 points)
Answer the following questions:

a. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with $X_i \sim N(\theta, 1)$. Let $\hat{\theta}$ be the mle of θ. Is $\hat{\theta}$ MVUE of θ?

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\theta)^2} \qquad L = (2\pi)^{-n/2} e^{-\frac{1}{2}\sum(x-\theta)^2}$$

$$\ln L = -\frac{n}{2}\ln 2\pi - \frac{1}{2}\sum(x-\theta)^2 \qquad \frac{\partial \ln L}{\partial \theta} = \sum(x-\theta) = 0 \Rightarrow \hat{\theta} = \bar{x}$$

$$var(\hat{\theta}) = \frac{1}{n}.$$

$$\frac{\partial^2 \ln}{\partial \theta^2} = -n \qquad \text{(CRAMER-RAO Low(er Bound))} \quad \frac{1}{-E(-n)} = \frac{1}{n} \checkmark \qquad \text{yep!}$$

b. If X has a binomial distribution with parameters n and p, then $\hat{p}_1 = \frac{X}{n}$ is an unbiased estimator of p. Another estimator of p is $\hat{p}_2 = \frac{X+1}{n+2}$. Find the bias of $\hat{p}_2$.

$$E\hat{p}_2 = E\frac{X+1}{n+2} = \frac{EX+1}{n+2} = \frac{np+1}{n+2}$$

$$B = E\hat{p}_2 - p = \frac{np+1}{n+2} - p = \frac{np+1-np-2p}{n+2}$$

$$B = \frac{1-2p}{n+2}$$

c. Refer to part (b). Find $MSE(\hat{p}_1)$ and $MSE(\hat{p}_2)$.

$$MSE(\hat{p}_1) = VAR(\hat{p}_1) = VAR\left(\frac{X}{n}\right) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}.$$

$$MSE(\hat{p}_2) = VAR(\hat{p}_2) + B^2 = VAR\left(\frac{X+1}{n+2}\right) + \left(\frac{1-2p}{n+2}\right)^2$$

$$= \frac{np(1-p)}{(n+2)^2} + \frac{(1-2p)^2}{(n+2)^2} = \frac{np - np^2 + 1 + 4p^2 - 4p}{(n+2)^2} \ldots$$

d. Assume that you have a sample mean $\bar{X}$ and a sample variance S^2 based on a random sample $X_1, X_2, \ldots, X_n$ from a normal distribution. Suppose you want to predict a new value X_p. Find the distribution of $X_p - \bar{X}$ and use it together with the distribution of S^2 to form a ratio that follows the t distribution. Note: X_p is independent of $X_1, X_2, \ldots, X_n$.

$$E(X_p - \bar{X}) = 0, \qquad VAR(X_p - \bar{X}) = \sigma^2 + \frac{\sigma^2}{n} = \sigma^2\left(1 + \frac{1}{n}\right)$$

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

$$\frac{\dfrac{X_p - \bar{X} - 0}{\sigma\sqrt{1+\frac{1}{n}}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2} \Big/ n-1}} = \frac{X_p - \bar{X}}{S\sqrt{1+\frac{1}{n}}} \sim t_{n-1}$$

Problem 4 **(25 points)**
Consider k independent random variables that have normal distributions with unknown means $\mu_1, \mu_2, \ldots, \mu_k$, respectively, and unknown but common variance σ^2. Random samples of size $n_i, i = 1, 2, \ldots, n_k$ are selected from these k populations. Answer the following questions:

a. Write the likelihood and log-likelihood function that involves all the information above.

$$L = (2\pi\sigma^2)^{-\frac{(n_1+\cdots+n_k)}{2}} e^{-\frac{1}{2}\sum\left(\frac{x_i-\mu_1}{\sigma}\right)^2} \cdots e^{-\frac{1}{2}\sum\left(\frac{x_i-\mu_k}{\sigma}\right)^2}$$

$$\ln L = -\frac{(n_1+\cdots+n_k)}{2}\ln 2\pi\sigma^2 - \frac{1}{2}\sum\left(\frac{x_i-\mu_1}{\sigma}\right)^2 - \cdots - \frac{1}{2}\sum\left(\frac{x_i-\mu_k}{\sigma}\right)^2$$

b. Use the method of maximum likelihood to find estimates of $\mu_i, i = 1, 2, \ldots, k$ and σ^2. Denote these estimates with $\hat{\mu}_i, i = 1, 2, \ldots, k$ and $\hat{\sigma}^2$.

$$\frac{\partial \ln L}{\partial \mu_i} = \sum\left(\frac{x_i-\mu_1}{\sigma^2}\right) = 0 \Rightarrow \mu_i = \bar{x}_i$$

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{(n_1+\cdots+n_k)}{2\sigma^2} + \frac{1}{2\sigma^4}\sum(x_i-\mu_1)^2 + \cdots + \frac{1}{2\sigma^4}\sum(x_i-\mu_k)^2 = 0$$

$$\hat{\sigma}^2 = \frac{\sum(x_i-\mu_1)^2 + \cdots + \sum(x_{ki}-\mu_k)^2}{n_1+\cdots+n_k}$$

Finally, $$\hat{\sigma}^2 = \frac{\sum(x_{1i}-\bar{x}_1)^2 + \cdots + \sum(x_{ki}-\bar{x}_k)^2}{n_1+\cdots+n_k}$$

c. Is $\hat{\sigma}^2$ unbiased estimator of σ^2? If not, give the unbiased estimator of σ^2 and denote it with S_k^2.

$$E\hat{\sigma}^2 = \frac{(n_1-1)\sigma^2 + \cdots + (n_k-1)\sigma^2}{n_1+\cdots+n_k} = \frac{(n_1+n_2+\cdots+n_k-k)\sigma^2}{n_1+\cdots+n_k}$$

Therefore the unbiased estimator will be.

$$S_k^2 = \frac{n_1+\cdots+n_k}{(n_1+n_2+\cdots+n_k-k)}\hat{\sigma}^2 = \frac{\sum(x_{1i}-\bar{x}_1)^2 + \cdots + \sum(x_{ki}-\bar{x}_k)^2}{n_1+n_2+\cdots+n_k-k}$$

d. Find the distribution associated with S_k^2.

$$\frac{(n_1+n_2+\cdots+n_k-k)S_k^2}{\sigma^2} \sim \chi^2_{n_1+n_2+\cdots+n_k-k}$$