University of California, Los Angeles
Department of Statistics

Statistics 100B **Instructor: Nicolas Christou**

Practice 5 - solutions

EXERCISE 1
Let $X_1, X_2, \cdots, X_n$ be independent and identically distributed random variables from a Poisson distribution with parameter λ. We know that the maximum likelihood estimate of λ is $\hat{\lambda} = \bar{x}$.

a. Find the variance of $\hat{\lambda}$.
Answer:

$$Var(\bar{X}) = Var\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{n\lambda}{n^2} = \frac{\lambda}{n}.$$

b. Is $\hat{\lambda}$ an MVUE?
Answer: We need to find the lower bound of the Cramer-Rao inequality:

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!} \Rightarrow lnf(x) = xln\lambda - \lambda - lnx!$$

Let's find the first and second derivatives w.r.t. λ.

$$\frac{\partial lnf(x)}{\partial \lambda} = \frac{x}{\lambda} - 1 \text{ and } \frac{\partial^2 lnf(x)}{\partial \lambda^2} = -\frac{x}{\lambda^2}.$$

Therefore,

$$\frac{1}{-nE\left(\frac{\partial^2 lnf(x)}{\partial \lambda^2}\right)} = \frac{1}{-nE(-\frac{X}{\lambda^2})} = \frac{\lambda^2}{\lambda n} = \frac{\lambda}{n}.$$

Therefore, $\hat{\lambda}$ is MVUE.

c. Is $\hat{\lambda}$ a consistent estimator of λ?
Answer:
It is consistent because it is unbiased and its variance is equal to zero as $n \to \infty$.
$E(\hat{\lambda}) = \lambda$ and

$$\lim_{n\to\infty} \frac{\lambda}{n} = 0$$

.

EXERCISE 2
Suppose that two independent random samples of n_1 and n_2 observations are selected from two normal populations. Further, assume that the populations possess a common variance σ^2 which is unknown. Let the sample variances be S_1^2 and S_2^2 for which $E(S_1^2) = \sigma^2$ and $E(S_2^2) = \sigma^2$.

a. Show that the pooled estimator of σ^2 that we derived in class below is unbiased.

$$S^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$$

Answer:
Both S_1^2 and S_2^2 are unbiased estimators of σ^2. Therefore:

$$E(S^2) = E\left(\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}\right) = \frac{(n_1-1)E(S_1^2) + (n_2-1)E(S_2^2)}{n_1+n_2-2} =$$

$$\frac{(n_1-1)\sigma^2 + (n_2-1)\sigma^2}{n_1+n_2-2} = \sigma^2.$$

b. Find the variance of S^2.
Answer:
We use the result that $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ as follows:

$$(n_1 + n_2 - 2)S^2 = (n_1 - 1)S_1^2 + (n_2 - 1)S_2^2$$

Or

$$\frac{(n_1 + n_2 - 2)S^2}{\sigma^2} = \frac{(n_1 - 1)S_1^2}{\sigma^2} + \frac{(n_2 - 1)S_2^2}{\sigma^2}$$

Therefore,

$$\frac{(n_1 + n_2 - 2)S^2}{\sigma^2} \sim \chi^2_{n_1+n_2-2}$$

Finally,

$$Var\left(\frac{(n_1 + n_2 - 2)S^2}{\sigma^2}\right) = 2(n_1 + n_2 - 2) \Rightarrow Var(S^2) = \frac{2(n_1 + n_2 - 2)\sigma^4}{(n_1 + n_2 - 2)^2} = \frac{2\sigma^4}{n_1 + n_2 - 2}.$$

EXERCISE 3

Suppose $Y_i = \beta_1 x_i + \epsilon_i$. In this simple regression model through the origin x_i is non-random, β_1 is unknown parameter, and $\epsilon_i \sim N(0, \sigma)$.

a. Find the mean of Y_i.

b. Find the variance of Y_i.

c. What distribution does Y_i follow? Write the pdf of Y_i.

d. Write down the likelihood function based on n observations of Y and x.

e. Find the maximum likelihood estimate of β_1. Denote it with $\hat{\beta}_1$.

f. Show that $\hat{\beta}_1$ is unbiased estimator of β_1.

g. Find the variance of $\hat{\beta}_1$.

h. What is the distribution of $\hat{\beta}_1$?

i. Find the maximum likelihood estimate of σ^2.

Solution:

Let $Y_i = \beta_1 x_i + \epsilon_i$, where $\epsilon_i \sim N(0, \sigma)$. The x_i's are non-random.

a. $E(Y_i) = \beta_1 x_i$.

b. $var(Y_i) = \sigma^2$

c. $Y_i \sim N(\beta_1 x_i, \sigma)$. And its pdf function is:

$$f(y_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{y_i - \beta_1 x_i}{\sigma})^2}$$

d.
$$L = (\sigma^2 2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{n}(\frac{y_i - \beta_1 x_i}{\sigma})^2}$$

e. From part (d) we can write the log-likelihood function:

$$ln(L) = -\frac{n}{2} ln(\sigma^2 2\pi) - \frac{1}{2}\sum_{i=1}^{n}(\frac{y_i - \beta_1 x_i}{\sigma})^2.$$

To find the estimate of β_1 we take the partial derivative of the log-likelihhod w.r.t. to β_1, set it equal to zero and solve:

$$\frac{\partial ln(L)}{\partial \beta_1} = \frac{2}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta_1 x_i)x_i = 0$$

Solving for β_1 we get:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}.$$

f. The previous estimate is unbiased because:

$$E(\hat{\beta}_1) = \frac{\sum_{i=1}^{n} x_i E(y_i)}{\sum_{i=1}^{n} x_i^2} = \beta_1.$$

g. The variance of this estimate is:

$$var(\hat{\beta}_1) = var(\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}) = \frac{\sigma^2}{\sum_{i=1}^{n} x_i^2}.$$

h. Distirbution of $\hat{\beta}_1$:

$$\hat{\beta}_1 \sim N(\beta_1, \frac{\sigma}{\sqrt{\sum_{i=1}^{n} x_i^2}}).$$

i. Maximum likelihood of σ^2:

$$\frac{\partial ln(L)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(y_i - \beta_1 x_i)^2 = 0 \Rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{\beta}_1 x_i)^2}{n}.$$

EXERCISE 4

a. We need to find first the mean of X:

$$E(X) = \int_{-1}^{1} x\frac{1+\theta x}{2}dx = [\frac{x^2}{4} + \frac{\theta x^3}{6}]_{-1}^{1} = \frac{\theta}{3}.$$

Therefore the method of moments estimator of θ is $\frac{\theta}{3} = \bar{x} \Rightarrow \hat{\theta} = 3\bar{x}$.

b. Yes, $\hat{\theta}$ is unbiased because $E(\hat{\theta}) = 3E(\bar{x}) = 3\mu = 3\frac{\theta}{3} = \theta$.

c. We need to find first the variance of X:

$$var(X) = EX^2 - (E(X))^2 = \int_{-1}^{1} x^2\frac{1+\theta x}{2}dx - (\frac{\theta}{3})^2 = [\frac{x^3}{6} + \frac{\theta x^4}{8}]_{-1}^{1} - (\frac{\theta}{3})^2 = \frac{1}{3} - (\frac{\theta}{3})^2 = \frac{3-\theta^2}{9}.$$

Thererfore, $var(\hat{\theta}) = var(3\bar{x}) = 9\frac{\sigma^2}{n} = 9\frac{3-\theta^2}{9n} = \frac{3-\theta^2}{n}$.

d. Yes, $\hat{\theta}$ is consistent because it is unbiased, and its variances goes to zero as n goes to infinity.

EXERCISE 5

a. Using the method of moments we get:

$$E(X) = \alpha\beta \Rightarrow \alpha\beta = \bar{x} \Rightarrow \beta = \frac{\bar{x}}{\alpha},$$

and

$$\sigma^2 = EX^2 - (E(X))^2 = \alpha\beta^2 \Rightarrow \frac{\sum_{i=1}^n x_i^2}{n} - \left(\frac{\sum_{i=1}^n x_i}{n}\right)^2 = \alpha\beta^2$$

From the two expressions above we get:

$$\begin{aligned}
\alpha\beta^2 &= \frac{\sum_{i=1}^n x_i^2}{n} - \left(\frac{\sum_{i=1}^n x_i}{n}\right)^2 \\
\alpha\frac{\bar{x}^2}{\alpha^2} &= \frac{\sum_{i=1}^n x_i^2}{n} - \left(\frac{\sum_{i=1}^n x_i}{n}\right)^2 \\
\hat{\alpha} &= \frac{\bar{x}^2}{\frac{\sum_{i=1}^n x_i^2}{n} - \left(\frac{\sum_{i=1}^n x_i}{n}\right)^2} = \frac{\bar{x}^2}{\frac{1}{n}\left(\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}\right)} \Rightarrow \hat{\alpha} = \frac{n\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2},
\end{aligned}$$

and $\hat{\beta} = \frac{\sum_{i=1}^n (x_i-\bar{x})^2}{n\bar{x}}$.

b. $E(X) = \frac{3\theta}{2} = \bar{x} \Rightarrow \hat{\theta} = \frac{2}{3}\bar{x}$.

c. The estimator of part (b) is consistent because:

1. It is unbiased: $E(\hat{\theta}) = \frac{2}{3}\mu = \frac{2}{3}\frac{3\theta}{2} = \theta$.
2. $var(\hat{\theta}) = \frac{4}{9}\frac{\sigma^2}{n} = \frac{4}{9}\frac{9\theta^2}{12n} = \frac{\theta^2}{3n} \to 0$ as $n \to \infty$.

EXERCISE 6
We need to find first the cdf of X:

$$F(X) = \int_{\theta}^{x} 3\theta^3 u^{-4} du = [-\frac{3\theta^3 u^{-3}}{3}]_{\theta}^{x} = 1 - \theta^3 x^{-3}.$$

a. Therefore the pdf of $X_{(1)}$ is:

$$g_1(x) = n\,(1 - F(x))^{n-1} f(x) = n\left(\theta^3 x^{-3}\right)^{n-1} 3\theta^3 x^{-4} = 3n\theta^{3n} x^{-3n-1}.$$

b. The mean of $X_{(1)}$:

$$EX_{(1)} = \int_{\theta}^{\infty} x 3n\theta^{3n} x^{-3n-1} dx = 3n\theta^{3n}[\frac{x^{-3n+1}}{-3n+1}]_{\theta}^{\infty} = \frac{3n\theta}{3n-1}.$$

c. $B = E\hat{\theta} - \theta = \frac{3n\theta}{3n-1} - \theta = \frac{\theta}{3n-1}$.

EXERCISE 7
Solution:

a. We must show that the sum of the residuals is equal to zero.

$$\sum_{i=1}^{n} e_i = \sum_{i=1}^{n}(y_i - \hat{y}_i) = \sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) =$$

$$\sum_{i=1}^{n}(y_i - (\bar{y} - \hat{\beta}_1\bar{x}) - \hat{\beta}_1 x_i) = \sum_{i=1}^{n}(y_i - \bar{y}) - \hat{\beta}_1 \sum_{i=1}^{n}(x_i - \bar{x}) = 0.$$

b. The $\text{cov}(\bar{Y}, \hat{\beta}_1) = 0$ because:

$$\begin{aligned}
\text{cov}(\bar{Y}, \hat{\beta}_1) &= Cov\left(\frac{1}{n}\sum_{i=1}^{n} Y_i, \frac{\sum_{i=1}^{n}(x_i - \bar{x})Y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right) \\
&= \frac{1}{n}\frac{1}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\left[(x_1 - \bar{x})Var(Y_1) + (x_2 - \bar{x})Var(Y_2) + \cdots + (x_n - \bar{x})Var(Y_n)\right] \\
&= \frac{1}{n}\frac{\sum_{i=1}^{n}(x_i - \bar{x})\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sigma^2}{n}\frac{\sum_{i=1}^{n}(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = 0.
\end{aligned}$$

Note: $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$ and $cov(Y_i, Y_j) = 0$ because the Y's are independent.

EXERCISE 8
Solution:

$$\begin{aligned}
cov(e_i, e_j) &= cov(Y_i - \hat{Y}_i, Y_j - \hat{Y}_j) = cov(Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i, Y_j - \hat{\beta}_0 - \hat{\beta}_1 x_j) \\
&= cov(Y_i - \bar{Y} + \hat{\beta}_1\bar{x} - \hat{\beta}_1 x_i, Y_j - \bar{Y} + \hat{\beta}_1\bar{x} - \hat{\beta}_1 x_j,) \\
&= cov\left(Y_i - \bar{Y} - \hat{\beta}_1(x_i - \bar{x}), Y_j - \bar{Y} - \hat{\beta}_1(x_j - \bar{x})\right) \\
&= cov(Y_i, Y_j) - cov(Y_i, \bar{Y}) - (x_j - \bar{x})cov(Y_i, \hat{\beta}_1) \\
&\quad -cov(\bar{Y}, Y_j) + var(\bar{Y}) + (x_j - \bar{x})cov(\bar{Y}, \hat{\beta}_1) \\
&\quad -(x_i - \bar{x})cov(\hat{\beta}_1, Y_j) + (x_i - \bar{x})cov(\hat{\beta}_1, \bar{Y}) + (x_i - \bar{x})(x_j - \bar{x})var(\hat{\beta}_1) \\
&= 0 - \frac{\sigma^2}{n} - \frac{(x_j - \bar{x})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sigma^2 \\
&\quad -\frac{\sigma^2}{n} + \frac{\sigma^2}{n} + 0 \\
&\quad -\frac{(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sigma^2 + 0 + \frac{(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sigma^2 \\
&= \sigma^2\left[-\frac{1}{n} - \frac{(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right].
\end{aligned}$$