LECTURE 11

Example 16.4. Let X_1, X_2,..., X_n denote a sample from a uniform $(0, \theta)$ distribution, where θ is assumed unknown. Since

$$E(X_i) = \frac{\theta}{2},$$

a "natural" estimator to consider is the unbiased estimator

$$d_1 = d_1(X_1, X_2, ..., X_n) = \frac{2\sum_{i=1}^n X_i}{n}.$$

Since

$$E(d_1) = E\left[\frac{2\sum_{i=1}^n X_i}{n}\right] = \frac{2}{n} E\left[\sum_{i=1}^n X_i\right] =$$

$$\frac{2}{n}\sum_{i=1}^n E[X_i] = \frac{2}{n}\sum_{i=1}^n \frac{\theta}{2} = \frac{n\cdot\theta}{n} = \theta,$$

it follows that

$$r(d_1, \theta) = E(d_1 - \theta)^2 = E(d_1 - E(d_1))^2 = \mathrm{Var}_\theta(d_1) = \mathrm{Var}_\theta\left[\frac{2\sum_{i=1}^n X_i}{n}\right] = \frac{4}{n^2}\mathrm{Var}_\theta\left(\sum_{i=1}^n X_i\right) =$$

$$= \frac{4}{n^2}\sum_{i=1}^n \mathrm{Var}_\theta(X_i) = \frac{4}{n^2}\sum_{i=1}^n \frac{\theta^2}{12} = \frac{4}{n}\frac{\theta^2}{12} = \frac{\theta^2}{3\,n}$$

A second possible estimator of θ is the maximal likelihood estimator, which is given by

$$d_2 = d_2(X_1, X_2, ..., X_n) = \max_{1\le i\le n} X_i = \max(X_1, X_2, ..., X_n).$$

To compute the mean square error of d_2 as an estimator of θ, we need to first compute its mean (so as to determine its bias) and variance. To do so, note that the distribution function of d_2 is as follows:

$$F_2(x) = F_{d_2}(x) = P\{d_2(X_1, X_2, ..., X_n) \le x\} = P\{\max_{1\le i\le n} X_i \le x\} =$$

$$= P\left(\omega : X_1 \le x \bigcap X_2 \le x \bigcap ... \bigcap X_n \le x\right) = \prod_{i=1}^n P\{X_i \le x\} =$$

$$= [P(X_1 \le x)]^n = \begin{cases} 0 & \text{if} \quad x < 0 \\ \left(\frac{x}{\theta}\right)^n & \text{if} \quad 0 \le x < \theta \\ 1 & \text{if} \quad x \ge \theta \end{cases}$$

Hence, upon differentiating, we obtain that the density function of d_2 is

$$f_2(x) = \begin{cases} 0 & \text{if} \quad x \notin (0, \theta) \\ \frac{n\, x^{n-1}}{\theta^n} & \text{if} \quad 0 \le x \le \theta \end{cases}$$

Therefore,

$$E(d_2) = \int_0^\theta x \frac{n\, x^{n-1}}{\theta^n}\, dx = \frac{n}{n+1}\theta. \tag{16.5}$$

Also

$$E(d_2^2) = \int_0^\theta x^2 \frac{n\, x^{n-1}}{\theta^n}\, dx = \frac{n}{n+2}\theta^2 \tag{16.6}$$

and so

$$\mathrm{Var}_\theta(d_2) = \frac{n}{n+2}\theta^2 - \left(\frac{n}{n+1}\theta\right)^2 =$$

$$= n\,\theta^2 \left[\frac{1}{n+2} - \frac{n}{(n+1)^2}\right] = \frac{n\,\theta^2}{(n+2)(n+1)^2}.$$

Hence

$$r(d_2, \theta) = E(d_2 - \theta)^2 = E\left[(d_2 - Ed_2) + (Ed_2 - \theta)\right]^2 =$$

$$= (E(d_2) - \theta)^2 + 2E[(Ed_2 - \theta)(d_2 - Ed_2)] + E[(d_2 - Ed_2)^2] =$$

$$= (E(d_2) - \theta)^2 + 2(Ed_2 - \theta)\, E[(d_2 - E(d_2))] + E[(d_2 - Ed_2)^2] = (E(d_2) - \theta)^2 + \mathrm{Var}_\theta(d_2) =$$

$$= \left[\frac{n}{n+1}\theta - \theta\right]^2 + \frac{n\,\theta^2}{(n+2)(n+1)^2} = \frac{\theta^2}{(n+1)^2} + \frac{n\,\theta^2}{(n+2)(n+1)^2} =$$

$$= \frac{\theta^2}{(n+1)^2}\left[1 + \frac{n}{n+2}\right] = \frac{2\theta^2}{(n+1)(n+2)}. \tag{16.7}$$

Since

$$\frac{2\theta^2}{(n+1)(n+2)} \le \frac{\theta^2}{3\,n}, \qquad n = 1, 2, ...$$

it follows that d_2 is more superior estimator of θ that is d_1.

Equation (16.5) suggests the use of even another estimator – namely, the unbiased estimator

$$(1 + 1/n) d_2(X_1, X_2, ..., X_n) = (1 + 1/n)\max(X_1, X_2, ..., X_n).$$

However, rather than considering this estimator directly, let us consider all estimators of the form

$$d_c(X_1, X_2, ..., X_n) = c\, \max(X_1, X_2, ..., X_n) = c\, d_2(X_1, X_2, ..., X_n),$$

where c is a given constant. The mean square error of this estimator is (we use (16.6) and (16.5))

$$r(d_c(X_1, X_2, ..., X_n, \theta) = E(d_c - \theta)^2 = \mathrm{Var}_\theta(d_c(X_1, X_2, ..., X_n)) + (E[d_c(X_1, X_2, ..., X_n)] - \theta)^2 =$$

$$= c^2 \mathrm{Var}_\theta(d_2(X_1, X_2, ..., X_n)) + (c\, E[d_2(X_1, X_2, ..., X_n)] - \theta)^2 =$$

$$= \frac{c^2\, n\, \theta^2}{(n+2)(n+1)^2} + \theta^2 \left(\frac{c\, n}{n+1} - 1 \right)^2 \tag{16.8}$$

To determine the constant c resulting in minimal mean square error, we differentiate to obtain

$$\frac{d}{dc} r(d_c(X_1, X_2, ..., X_n, \theta) = \frac{2c\, n\, \theta^2}{(n+2)(n+1)^2} + \frac{2\theta^2 n}{n+1} \left(\frac{c\, n}{n+1} - 1 \right).$$

Equating this to 0 shows that the best constant c – call it $\widehat{c}$ – is such that

$$\frac{\widehat{c}}{n+2} + \widehat{c}\, n - (n+1) = 0$$

or

$$\widehat{c} = \frac{(n+1)(n+2)}{n^2 + 2n + 1} = \frac{n+2}{n+1}.$$

Substitution this value of c into equation (16.8) yields that

$$r\left(\frac{n+2}{n+1} \max(X_1, X_2, ..., X_n), \theta \right) = \frac{n\, (n+2)\, \theta^2}{(n+1)^4} + \theta^2 \left(\frac{n(n+2)}{(n+1)^2} - 1 \right)^2 =$$

$$= \frac{n\, (n+2)\, \theta^2}{(n+1)^4} + \frac{\theta^2}{(n+1)^4} = \frac{\theta^2}{(n+1)^2}.$$

A comparison with equation (16.7) shows that the (biased) estimator

$$\frac{n+2}{n+1} \max(X_1, X_2, ..., X_n)$$

has about half the mean square error of the maximum likelihood estimator $\max(X_1, X_2, ..., X_n)$.

APPENDIX. THE EXPONENTIAL RANDOM VARIABLES.

Let us calculate $E\eta$ and $\mathrm{Var}\,(\eta)$. By definition we have

$$E\eta = \int_{-\infty}^{+\infty} x\, f_\eta(x)\, dx = \int_0^{+\infty} x\,\lambda\, e^{-\lambda\, x}\, dx.$$

Integrating by parts $\left(\lambda e^{-\lambda\, x}\, dx = dv, u = x\right)$ yields

$$E\eta = -\, xe^{-\lambda\, x}\Big|_0^{+\infty} + \int_0^{+\infty} e^{-\lambda\, x}\, dx = 0 - \frac{1}{\lambda} e^{-\lambda\, x}\Big|_0^{+\infty} = \frac{1}{\lambda}.$$

To obtain the variance of η, we first find $E\left[\eta^2\right]$.

$$E\left[\eta^2\right] = \int_0^{+\infty} x^2\,\lambda\, e^{-\lambda\, x}\, dx.$$

Integrating by parts $\left(\lambda e^{-\lambda\, x}\, dx = dv, u = x^2\right)$ gives

$$E\left[\eta^2\right] = -\, x^2\, e^{-\lambda\, x}\Big|_0^{+\infty} + 2\int_0^{+\infty} x\, e^{-\lambda\, x}\, dx = 0 + \frac{2}{\lambda}\, E\eta = \frac{2}{\lambda^2}.$$

Hence

$$\mathrm{Var}\,(\eta) = E\left[\eta^2\right] - \left[E\eta\right]^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}.$$

Thus the mean of the exponential distribution is the reciprocal of its parameter λ and the variance is the mean squared.

The key property of an exponential random variable is that it is memoryless, where we say that a nonnegative random variable $\eta(\omega)$ is memoryless if

$$P(\omega:\ \ \eta(\omega) > s + t/\eta(\omega) > t) = P(\omega : \eta(\omega) > s), \quad \text{for all} \quad s, t \geq 0. \tag{17.5}$$

To understand why equation (17.5) is called the memoryless property, imagine that η represents the length of time that a certain item functions before failing. Now let us consider the probability that an item that is still functioning at age t will continue to function for at least an additional time s. Since this will be the case if the total functional lifetime of the item exceeds $t + s$ given that the item is still functioning at t, we see that

$$P(\text{additional functional life of } t\text{-unit-old item exceeds}\, s) = P(\omega:\ \ \eta(\omega) > s + t/\eta(\omega) > t).$$

Thus, we see that equation (17.5) states that the distribution of additional functional life of an item of age t is the same as that of a new item. In other words, when equation (17.5) is satisfied, there is no need to remember the age of a functional item since as long as it is functional it is "as good as new".

The condition in equation (17.5) is equivalent to (using definition of conditional probability)

$$\frac{P(\omega: \ \eta(\omega) > s + t \cap \eta(\omega) > t)}{P(\omega: \eta(\omega) > t)} = P(\omega : \eta(\omega) > s).$$

or

$$P(\omega: \ \eta(\omega) > s + t) = P(\omega : \eta(\omega) > s)\, P(\omega : \eta(\omega) > t). \tag{17.6}$$

When $\eta(\omega)$ is an exponential random variable, then

$$P(\omega : \eta(\omega) > x) = \exp(-\lambda\, x), \quad \text{if} \quad x > 0$$

and so equation (17.6) is satisfied, because

$$\exp(-\lambda(t + s)) = \exp(-\lambda s)\, \exp(-\lambda t).$$

Hence, exponentially distributed random variables are memoryless. It fact it can be shown that they are the only continuous random variables that are memoryless:

$$f(x + y) = f(x)\, f(y). \tag{17.7}$$

The only continuous function defined on the whole real axis and satisfying the equation (17.7) is an exponential function (except for the function identically equal to zero).

Among discrete random variables only geometrical distributed random variable is memoryless:

$$P(\eta = n) = (1 - p)^n p, \qquad n = 0, 1, 2, ...$$

The following lemma presents another property of the exponential distribution.

Lemma. If η_1, η_2,..., η_n are independent exponential random variables having respective parameters λ_1, λ_2,..., λ_n, then $\min(\eta_1, \eta_2, ..., \eta_n)$ is exponential with parameter $\sum\limits_{i=1}^{n} \lambda_i$.

Proof. Since the smallest value of a set of numbers is greater than x if and only if all values are greater than x, we have (using independence)

$$P(\min(\eta_1, \eta_2, ..., \eta_n) > x) = P(\eta_1 > x \cap \eta_2 > x, ..., \cap \eta_n > x) =$$

$$= \prod_{i=1}^{n} P(\eta_i > x) = \prod_{i=1}^{n} \exp(-\lambda_i x) = \exp^{\left(-\sum_{i=1}^{n} \lambda_i x\right)}, \quad \text{for} \quad x \geq 0$$

Therefore

$$P(\min(\eta_1, \eta_2, ..., \eta_n) \leq x) = 1 - \exp^{\left(-\sum_{i=1}^{n} \lambda_i x\right)}, \quad \text{for} \quad x \geq 0$$

and 0 for $x \leq 0$.

The parameter λ is called the rate of the exponential distribution.

Example 16.5. Suppose that the amount of time one spends in a bank is exponentially distributed with mean ten minutes, that is $\lambda = \frac{1}{10} = 0.1$.

a) What is the probability that a customer will spend more than fifteen minutes in the bank?

b) What is the probability that a customer will spend more than fifteen minutes in the bank given that he is still in the bank after ten minutes?

Solution. If $\eta(\omega)$ represents the amount of time that the customer spends in the bank, then the first probability (case a)) is just

$$P(\eta > 15) = \exp(-15\,\lambda) = e^{-3/2} \approx 0.22$$

The second question asks for the probability that a customer who has spent ten minutes in the bank will have to spend at least five more minutes. However, since exponential distributions are memoryless, this must equal the probability that an entering customer spends at least five minutes in the bank. That is, the desired probability is just

$$P(\eta > 5) = \exp(-5\,\lambda) = e^{-1/2} \approx 0.604.$$