Politecnico di Milano
STATISTICS (079086), 2009/2010, Prof. A.Barchielli
Problem set n. 3

15th April 2010

Exercise 1 Let (X_1, X_2, X_3, X_4) be a random sample from a Uniform distribution over the interval$(\theta-1, \theta+1)$, with θ unknown.

(a) Compute the mean and variance of X_i. *[$E(X_i) = \theta; V(X_i) = 1/3$]*

(b) Set
$$Y = \frac{X_1 + 3X_2}{8} + \frac{X_3 + X_4}{4}.$$
Compute the mean and variance of Y. *[$E(Y) = \theta; V(Y) = 3/32$]*

(c) Set
$$Z = \frac{X_1 + X_2 + X_3 + X_4}{4}.$$
Compute the mean and variance of Z. *[$E(Z) = \theta; V(Z) = 1/12$]*

(d) Which one, between Y and Z, is the best estimator of θ? *[Z].*

Exercise 2 Let $(X_1, \ldots, X_n)$ be a random sample from a discrete distribution such that
$$p(-1) = \frac{\theta}{2}, \quad p(0) = 1 - \theta, \quad p(1) = \frac{\theta}{2}$$

(a) For which values of θ the function p is a density? *[We need $0 \leq p_i \leq 1$ and $\sum p_i = 1$; thus, $0 \leq \theta \leq 1$]*

(b) Compute the mean and variance of X_1. *[$E(X) = 0; V(X) = \theta$]*

Consider two estimators of θ, defined respectively as $G_n = \frac{1}{n}\sum_{i=1}^{n} |X_i|$ and $T_n = |X_n|$.

(c) Are G_n and T_n unbiased? *[The estimator U_n of θ is said to be unbiased if $E(U_n) = \theta$. Both G_n and T_n are unbiased]*

(d) Compute the MSE of the two estimators G_n and T_n and say what happens if $n \to \infty$. *[$MSE_{U_n} = Var(U_n) + \left[E(U_n) - \theta\right]^2$; hence, $MSE_{G_n} = \frac{\theta(1-\theta)}{n}$ and $MSE_{T_n} = \theta(1-\theta)$. Note that $MSE_{G_n} \to 0$ for $n \to \infty$]*

(e) Which estimator is better? Why? *[G_n is better since its MSE is smaller for any $n > 1$ and moreover goes to 0 for $n \to \infty$]*

Exercise 3 X_1 is a sample from a Poisson distribution with parameter λ. Consider the following estimators for λ: $T_1(X_1) = X_1$ and $T_2(X_1) = 1$. Which one is the best according to the MSE? *[T_2 is better than $T1$ for $0.381 \leq \lambda \leq 2.618$; for all other values of λ, T_1 is better]*

Exercise 4 Let (X_1, X_2, X_3) be a random sample from a Bernoulli distribution. The following estimator has been proposed to estimate the mean μ of the distribution:

$$T = \frac{1}{12}X_1 + 3(X_2 + X_3).$$

a) Shows that T is biased and compute its MSE. *[$MSE_T = Var(T) + \left[E(U_n) - \theta\right]^2 = (\frac{1}{12^2} + 9 + 9)Var(X_1) + (\frac{73}{12}\mu - \mu)^2 = 18.0069\mu(1-\mu) + 25.84\mu^2$]*

b) Compute the constant c such that $W = cT$ is an unbiased estimator of μ, and compute its MSE. *[$c = \frac{12}{73}$; $MSE_W = \left(\frac{12}{73}\right)^2 18.0069\mu(1-\mu)$]*

c) Which of the two estimators is the best? *[W]*

Exercise 5 The number of people entering a Computer Shop in the center of Milan every day is a Poisson random variable with parameter λ. If during 30 days the total number of people who have entered the shop have been 1050, which is the maximum likelihood estimate of λ?
Find also the maximum likelihood estimator of $1/\lambda$; is it unbiased?

Exercise 6 Let $(X_1, \ldots, X_n)$ be a random sample from a Gaussian population with mean μ and variance $\sigma^2 = 16$.

a) Knowing that $n = 9$ and $\overline{x} = 5$, compute the confidence interval for μ at level 90%. *[The confidence interval for μ is given by: $(\overline{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}; \overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}})$. Using the data above we have the interval (2.807; 7.193)]*

b) Compute the confidence interval for μ at level 95%. *[(2.387; 7.613)]*

c) Determine the smaller sample size n such that the confidence interval for μ at level 95% is not longer than 4. *[The length of the confidence interval is given by $2z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$. With the data above, we need $n \geq 16$]*

Exercise 7** Let $X_1, \ldots, X_n$ be i.i.d. $\sim f(x, \theta) = \frac{1}{\theta}e^{-x/\theta}\mathbf{1}_{(0,\infty)}(x)$, where $\theta > 0$.

1. Find the density of the sample mean $\bar{X}$ [$\bar{X} \sim Gamma(n, \theta/n)$];
2. Find the density of the random variable $\frac{2n\bar{X}}{\theta}$ [$\frac{2n\bar{X}}{\theta} \sim Gamma(n, 2) \equiv \chi^2(2n)$];
3. Let $\alpha = 5\%$, $n = 3$ and $\theta = 2$: determine k such that $P(\bar{X} \leq k) = \alpha$ [Use the fact that $\frac{2n\bar{X}}{\theta}$, from the tables we have $k \simeq 1.64/3$]
4. Let $n = 3$ and $\theta = 1.49$: calculate $P(\bar{X} > 1.64/3)$
   [$P(\chi^2_6 > 2.2013) = 1 - F_{\chi^2_6}(2.2013) \simeq 1 - 0.1 = 0.9$]