Chapter VI: Introduction to Statistic Inference

PROBLEMS

Proposed Problems

1. The random variable X_1 is distributed according to a Normal $N(\mu, \sigma^2)$, and the random variable X_2, independent form the previous one, with distribution $N(2\mu, 3\sigma^2)$. We take a sample of size n_1 of the former variable and a sample of size n_2 of the latter. To estimate the value of the parameter μ we use the estimator $\hat{\mu} = a\,\bar{x}_1 + b\,\bar{x}_2$ where a and b are constants and $\bar{x}_1$ and $\bar{x}_2$are the sample means of the two samples. What conditions should the constant a and b satisfy to have that the estimator $\hat{\mu}$ is unbiased?

   SOLUTION:

   The distributions for the sample means are $\bar{x}_1 \sim N\left(\mu, \frac{\sigma^2}{n_1}\right)$ and $\bar{x}_2 \sim N\left(2\mu, \frac{3\sigma^2}{n_2}\right)$ therefore

   $$E[\hat{\mu}] = E[a\,\bar{x}_1 + b\,\bar{x}_2] = a\,E[\,\bar{x}_1] + b\,E[\,\bar{x}_2] = (a + 2b)\mu$$

   and to have a null bias the two constants have to satisfy the relation $(a + 2b) = 1$.

2. Given a sample of size $n = 3$ taken from a Normal random variable X with unknown mean μ and known variance $\sigma^2 = 1$, we consider the following estimator for the parameter μ

   $$\hat{\mu}_1 = \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3;$$
   $$\hat{\mu}_2 = \frac{1}{4}X_1 + \frac{1}{2}X_2 + \frac{1}{4}X_3;$$
   $$\hat{\mu}_3 = \frac{1}{8}X_1 + \frac{3}{8}X_2 + \frac{1}{2}X_3;$$

   where X_1, X_2 and X_3 are the observations. Check that the three estimators are unbiased and study their MSQ (Mean Square Error) . (Sept '99)

   SOLUTION:

   $E[\hat{\mu}_1] = \frac{1}{3}E[X_1] + \frac{1}{3}E[X_2] + \frac{1}{3}E[X_3] = \mu;$

   $E[\hat{\mu}_2] = \frac{1}{4}E[X_1] + \frac{1}{2}E[X_2] + \frac{1}{4}E[X_3] = \mu;$

   $E[\hat{\mu}_3] = \frac{1}{8}E[X_1] + \frac{3}{8}E[X_2] + \frac{1}{2}E[X_3] = \mu;$

   Since they are unbiased the MSE is equal to their variances and we have

   $Var[\hat{\mu}_1] = \frac{1}{9}Var[X_1] + \frac{1}{9}Var[X_2] + \frac{1}{9}Var[X_3] = \frac{1}{9}\sigma^2 = \frac{1}{9}$=0.333;

   $Var[\hat{\mu}_2] = \frac{1}{16}Var[X_1] + \frac{1}{4}Var[X_2] + \frac{1}{16}Var[X_3] = \frac{3}{8}\sigma^2 = \frac{3}{8}$=0.375;

   $Var[\hat{\mu}_3] = \frac{1}{64}Var[X_1] + \frac{9}{64}Var[X_2] + \frac{1}{4}Var[X_3] = \frac{26}{64}\sigma^2 = \frac{13}{12} = 0.406;$

   The most efficient estimator is $\hat{\mu}_1$.

3. We want to estimate the mean μ of a random variable X . To do this we measure 10 samples and we compute the sample mean $\bar{X}$ and their variance s_X^2. Say if the following sentences are true or false
   a) For the Central Limit Theorem we know that μ is a Normal random variable.
   b) $\bar{X}$ is an estimator for X.
   c) If we want to estimate μ we need at least 50 samples
   d) The sample mean given a sample set of data is a number and not a random variable.
   e) The sample mean $\bar{X}$ has always a Normal distribution
   f) If X is Normal then $\bar{X}$ is always Normal.
   g) To lower the variance of $\bar{X}$ to half of it, we have to take at least 100 samples
   h) Since the sample mean is an unbiased estimator, we know for sure that $\bar{X} = \mu$.

4. An analyst takes 100 samples of a Normal random variable X, with the aim to estimate its mean value μ. With this 100 samples he has two options to estimate μ:

   Estimator 1: Compute the sample mean of the 100 data

Estimator 2: Since$E[\bar{X}] = \mu$, a second option is to divide all data in 10 sets each of size 10, then to compute the sample means of all these sets $(\bar{X}_1, \bar{X}_2, \cdots, \bar{X}_{10})$ and finally to compute the average of these ten numbers.

Find:

a) Which options gives the better estimator μ?

b) How will the histogram of the 10 sample means used for the Estimator 2 look like?

5. The lifetime of a system up to a failure can be modeled by an exponential random variable T ~ Exp (λ). During some time we take notes of the duration of the system between failures and we get the following data measured in hours: 18, 94, 22, 143, 114. Estimate the parameter λ of the exponential distribution using the Moments' Method.

SOLUTION:

Since T is an exponential random variable we have $\lambda = 1/E[T]$. Therefore the Moments' Method we get $\hat{\lambda} = 1/\bar{T}$.

We can compute $\bar{T} = \frac{18+94+22+143+114}{5} = 78.2\ hours$ and therefore an estimation for λ is given by $\hat{\lambda} = \frac{1}{78.2} = 0.013 \frac{fault}{hour}$.