Politecnico di Milano
STATISTICS (079086), 2009/2010, Prof. A.Barchielli
Problem set n. 4

22th April 2010

Exercise 1 [Examination test Profs. Barchielli, Epifani, 06-03-08]

Let

$$10.42,\ 6.69,\ 11.18,\ 7.94,\ 9.65,\ 8.38,\ 11.80,\ 8.28,\ 10.84,\ 10.72$$

be the observed values of a random sample $X_1, \ldots, X_{10}$ from a Gaussian population having variance 2.25 and unknown mean μ.

1. Determine a bilateral confidence interval at level $\gamma = 90\%$ for the mean μ. [The sample mean is $\bar{x} = 9.59$ and the confidence interval is $(8.81, 10.37)$]

You can now enlarge the sample.

2. What is the minimum sample size necessary to be at least 90% confident that the error made in estimating μ by the sample mean is not greater than 0.26? [91]

Let n^* be the number found in 2.. You decide to enlarge the initial sample by extracting $n^* - 10$ observations $X_{11}, \ldots, X_{n^*}$. Suppose that the sample mean of the new $n^* - 10$ observations is 8.96.

3. Update the point estimate of μ using all information in the new enlarged sample $X_1, \ldots, X_{n^*}$ [9.03]; determine a bilateral confidence interval at level $\gamma = 90\%$ for the mean μ by using what you did in point 2.. [$(8.77, 9.29)$]

Exercise 2 The concentration of the enzyme GYH in the blood of 20 individuals has been measured obtaining the values $\bar{x} = 1.23$ and $s^2 = 0.4$ for the sample mean and for the sample variance, respectively.

1. Assuming normality, determine a bilateral confidence interval at level 95% for the mean [$(0.935; 1.525)$]
2. Suppose that the concentration of the enzyme is normally distributed with known variance $\sigma^2 = 0.4$, and calculate again a bilateral confidence interval at level 95% for the mean [$(0.953; 1.507)$]
3. Between the intervals found in 1. and 2., which one is the longest? Give an explanation.

Exercise 3** The systolic pressure of the blood of 90 healthy adults has been measured, giving the values 128.9 (mm. of mercury) and 17 (mm. of mercury) for the sample mean and for the standard deviation, respectively. Assume normality of the distribution.

1) Determine a bilateral confidence interval at level 95% for the mean pressure; [$(125.33, 132.47)$]
2) find the length of the interval; [7.14]
3) if the confidence level is raised up to 99%, does the length of the interval increase or decrease? Why?
4) if the values of the sample mean and the standard deviation were relative to a random sample of size 360, how long would the 95% confidence interval be? [$\simeq 3.51$]

Exercise 4 The quantity of sugar included in a pear juice is normally distributed. A random sample of $n = 10$ pear juices has sample variance $s^2 = 23.04\ mg^2$. Compute the bilateral confidence interval for σ^2 at level 98%. *[We know that* $1 - \alpha = P\big(\chi^2_{1-\frac{\alpha}{2},n-1} \leq (n-1)\frac{S^2}{\sigma^2} \leq \chi^2_{\frac{\alpha}{2},n-1}\big) = P\Big((n-1)S^2\frac{1}{\chi^2_{\frac{\alpha}{2},n-1}} \leq \sigma^2 \leq (n-1)S^2\frac{1}{\chi^2_{1-\frac{\alpha}{2},n-1}}\Big)$*; thus, the confidence interval is given by* $\Big((n-1)s^2\frac{1}{\chi^2_{\frac{\alpha}{2},n-1}}; (n-1)s^2\frac{1}{\chi^2_{1-\frac{\alpha}{2},n-1}}\Big) = (9.57; 99.32)$ *(*$\chi^2_{0.99,9} = 2.088$, $\chi^2_{0.01,9} = 21.666$*)]*

Exercise 5 Consider the family of densities given by

$$f(x;\vartheta)=\begin{cases}\frac{2x}{\vartheta} & \text{if } 0<x\le\vartheta\\ \frac{2(1-x)}{1-\vartheta} & \text{if } \vartheta<x<1\\ 0 & \text{elsewhere}\end{cases}$$

where $0<\vartheta<1$ is an unknown parameter.

1. Compute the mean and variance of $f\,(x;\vartheta)$ for any $\vartheta\in(0,1)$.
2. Given a random sample $X_1,\dots X_n$ extracted from a population with density $f\,(x;\vartheta)$, determine an estimator T_n of ϑ using the method of moments.
3. Verify that the sequence of estimators $(T_n)_n$ is unbiased and square mean consistent. Moreover, give the asymptotic distribution of $(T_n)_n$. *Justify all your answers.*
4. Construct an asymptotic bilateral confidence interval of approximate level 90% for θ, if $n=169$ and $\sum_{j=1}^{169}x_j=35.0$.

Solution 1.

$$\mathrm{E}(X)=\int_{-\infty}^{+\infty}xf\,(x;\vartheta)\;dx=\int_0^{\vartheta}x\frac{2x}{\vartheta}dx+\int_{\vartheta}^{1}x\frac{2\,(1-x)}{1-\vartheta}\;dx=\frac{\vartheta+1}{3}$$

$$\mathrm{E}(X^2)=\int_{-\infty}^{+\infty}x^2f\,(x;\vartheta)\;dx=\int_0^{\vartheta}x^2\frac{2x}{\vartheta}dx+\int_{\vartheta}^{1}x^2\frac{2\,(1-x)}{1-\vartheta}\;dx=\frac{\vartheta^2+\vartheta+1}{6}$$

$$\mathrm{Var}(X)=\mathrm{E}(X^2)-(\mathrm{E}(X))^2=\frac{\vartheta^2+\vartheta+1}{6}-\frac{(\vartheta+1)^2}{3^2}=\frac{\vartheta^2-\vartheta+1}{18}$$

2. Since $\mathrm{E}(X)=(\vartheta+1)/3$ the method of moments estimator is $T_n=3\overline{X}_n-1$.
3. T_n is unbiased by construction. It is square mean consistent since it is unbiased and

$$\mathrm{Var}(T_n)=\mathrm{Var}(3\overline{X}_n-1)=9\,\mathrm{Var}(\overline{X}_n)=\frac{9\,\mathrm{Var}(X_1)}{n}=\frac{(\vartheta^2-\vartheta+1)}{2n}\to 0,\ \text{ for } n\to\infty.$$

Finally, thanks to the central limit theorem, the asymptotic distribution of T_n is Gaussian with mean $\mathrm{E}(T_n)=\vartheta$ and variance $\mathrm{Var}(T_n)=(\vartheta^2-\vartheta+1)/(2n)$.

4. The asymptotic distribution of $(T_n-\vartheta)/\sqrt{(\vartheta^2-\vartheta+1)/(2n)}$ is $\mathcal{N}(0,1)$. It follows that an asymptotic bilateral confidence interval for ϑ of approximate level 90% is given by $T_n\pm 1.645\times\sqrt{(T_n^2-T_n+1)}/\sqrt{n}$. With the given data, $T_{169}=35/169\simeq 0.207$ and we are approximately 90%-confident that $0.1007<\vartheta<0.3133$.