MA 486-4A (Statistics), Dr Chernov — Midterm test #1

Show your work. — Thu, Feb 10, 2005

1. (15 pts) The following are the annual incomes (in thousands of dollars) of 10 randomly selected former graduates of a college:

23 27 35 32 29 34 28 33 96 23

(a) Determine the mode, the median, the quartiles and the IQR.
(b) Find the sample mean, the sample variance and the sample standard deviation.
(c) Which statistic – the sample mean, the median, or the mode – seems to better represent a typical annual income? Comment.

Answers:

(a) the mode is 23, the median is 30.5, the quartiles are 27 and 34.

(b) the sample mean is 36, the sample variance is 462, the sample standard deviation is 21.5

(c) the median

2. (10 pts) Let p denote the proportion of defective items manufactured in a production line. A quality test of randomly selected 50 items shows that 7 of them are defective.
(a) Construct an 80% confidence interval for p.
(b) Give the upper endpoint of a one-sided 80% confidence interval for p.

Answers:
(a) $0.14 \pm 1.282\sqrt{.14 \times 0.86/50} = 0.14 \pm 0.0629$. Hence, [0.077,0203].
(b) $0.14 + 0.842\sqrt{.14 \times 0.86/50} = 0.1813$

3. (20 pts) A random sample of size $n = 40$ from $N(\mu, \sigma^2)$ yielded $s^2 = 15$.
(a) Construct a 90% confidence interval for σ^2 by using percentiles of the χ^2 distribution.
(b) Construct the shortest 90% confidence interval for σ^2.
(c) Assume now that the sample has size $n = 400$ instead of $n = 40$. Construct a 90% confidence interval for σ^2.
[Bonus] Does the length of the confidence interval converge to zero as $n \to \infty$?

Answers:

(a) The values of $\chi^2_\alpha(39)$ are no available in the book, so use normal approximation:

$$\frac{15\sqrt{39}}{\sqrt{39} + 1.645\sqrt{2}} < \sigma^2 < \frac{15\sqrt{39}}{\sqrt{39} - 1.645\sqrt{2}}$$

so [10.93,23.90].

(b) Same, since $n > 30$.

(c) by using normal approximation

$$\frac{15\sqrt{399}}{\sqrt{399} + 1.645\sqrt{2}} < \sigma^2 < \frac{15\sqrt{399}}{\sqrt{399} - 1.645\sqrt{2}}$$

so [13.435,16.977].

[Bonus] yes

4. (25 pts) Let $x_1, \ldots, x_n$ be a random sample from the distribution with probability density function

$$f(x;\theta) = \frac{\theta - 1}{(x+1)^\theta} \quad \text{for } \ 0 < x < \infty$$

where $\theta > 1$ is an unknown parameter.
(a) Find the maximum likelihood estimate for θ.
(b) Find sufficient statistics, and make sure that the MLE is a function of them.
(c) Find the Rao-Cramer lower bound on the variance of unbiased estimates of θ.
(d) What is the asymptotic distribution of the MLE?

Answers:
(a) The MLE is

$$\hat{\theta} = \frac{n}{\sum \ln(x_i + 1)} + 1$$

(b) The only sufficient statistic is $\prod(x_i + 1)$. Its logarithm is $\sum \ln(x_i + 1)$
(c) The Rao-Cramer lower bound is

$$\mathrm{Var}(\hat{\theta}) \geq \sigma_n^2 = \frac{(\theta - 1)^2}{n}$$

(d) it is $N(\theta, \sigma_n^2)$.

5. (10 pts) A student believes that he has a loaded (unbalanced) coin, because when he flips it it seems to land on heads more frequently than tails. He decides to empirically estimate the probability p for the coin to land on heads by flipping it many times. How many times does he need to flip the coin in order to construct a 99% confidence interval with half-length 0.01?

Answer: 16590.

6. (20 pts) Let $x_1, \ldots, x_9$ and $y_1, \ldots, y_{13}$ be two independent random samples from distributions $N(\mu_x, \sigma^2)$ and $N(\mu_y, \sigma^2)$, respectively. Their sample means are $\bar{x} = 12.6$ and $\bar{y} = 6.1$. Their sample standard deviations are $s_x = 5.0$ and $s_y = 5.0$, respectively. Give the lower endpoint of a one-sided 95% confidence interval for the difference $\mu_x - \mu_y$. (Make sure that you are using the right assumptions!)

Answer: here the variances are equal (both are denoted by σ^2). Hence the the lower endpoint of the CI is

$$12.6 - 6.1 - t_{0.05}(20)\sqrt{\frac{8 \times 25 + 12 \times 25}{20}}\sqrt{\frac{1}{9} + \frac{1}{13}} = 2.76$$