Harald Schmidbauer
Sema Bayraktar
Yelda Yücel

İstanbul Bilgi Üniversitesi

Bus 274: Further Statistics for Business

Spring 2010

PROBLEM SHEET # 5

Problem 1: A public utility company makes an effort to investigate the annual water consumption (measured in cubic meters) of households with young families with children. A sample of annual water consumption in 10 households was:

$$103, 100, 178, 84, 115, 37, 228, 93, 75, 150$$

(It is $\bar{x} = 116.3$, $s^2 = 3049.3$.)

a) It is known, from a previous large-scale study, that the average annual water consumption of households without children is 100 cubic meters. Does the sample provide evidence that average consumption is higher for young families with children? (Hint: Assume that annual water consumption is normally distributed. Test a suitable null hypothesis.)

b) Build a 95% confidence interval for the standard deviation of annual water consumption. (Hint: Build a confidence interval for the *variance* first.)

c) Why are we — generally — interested in the standard deviation?

Problem 2: The audience rating of a certain TV program, which is broadcast weekly, was estimated twice: last week and this week, on the basis of information on 5000 randomly selected people in the relevant target group. The estimated rating was last week: 12.5%, this week: 13.7%. A comment on the website of the TV channel was: "The audience rating has increased dramatically!" Is this comment warranted by the data?

Problem 3: Which of the following statements is true?

A: The p-value is the probability of obtaining the calculated test statistic, and values beyond it, under the null hypothesis.
B: The p-value measures the probability that the null hypothesis is true.
C: The p-value measures the probability of making a type II error.
D: The larger the p-value, the stronger the evidence against the null hypothesis.
E: A large p-value indicates that the data is consistent with the alternative hypothesis.

Problem 4: Let $X_1, \ldots, X_n \sim \mathrm{N}(\mu, \sigma)$ be independent. If μ and σ^2 are unknown:

A: An unbiased estimator for σ^2 is given by $\frac{1}{\sqrt{n}} \sum (\bar{X} - X_i)^2$.
B: It holds that $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim t_{n-1}$.
C: It holds that $\frac{1}{\sigma^2} \sum (X_i - \bar{X})^2 \sim \chi^2_{n-1}$.
D: σ^2 can be used as an estimator for the sample variance.
E: A confidence interval for σ^2 will always be symmetric around the point estimate s^2.