

## Bus 274: Further Statistics for Business

Spring 2010

### PROBLEM SHEET # 2

**Problem 1:** A consumer protection agency monitors vegetable quantity in packages sold by a retailer. A sample of 10 packages of potatoes labeled “2500g” had weights (in grams)

2473, 2515, 2503, 2421, 2466, 2475, 2415, 2373, 2412, 2435

(The sample mean is  $\bar{x} = 2448.8$ ; the sample standard deviation is  $s = 44.84$ .)

- Assuming vegetable weight is normally distributed, compute a 95% confidence interval for the expected weight in a package.
- According to your result in (a): Should the consumer protection agency undertake further investigations at this point, or even take action against the retailer? Give reasons for your answer.

**Problem 2:** A newspaper wrote:

“The approval rate of politician A. among registered voters is 57%, with a margin of error of plus or minus four percentage points.”

- What does this mean? What is the margin of error?
- How many voters might have been in the random sample of this poll?
- According to this poll, can A. still be confident she has the support of a majority?

**Problem 3:** The data in file `supermarket.xls` show 508 observations of three variables concerning customers of a supermarket: total expenditure (in euros), day of the week, sex.

- Compute an approximate 95% confidence interval for the expectation of the random variable *total expenditure*.
- Compute an approximate 95% confidence interval for the kurtosis. (Hint: The standard error of the kurtosis is about 0.75.) Does the distribution of customer expenditure have a heavy tail, or even two heavy tails? What does this mean w.r.t. buying behaviour of customers in general?

**Problem 4:** Confidence intervals for the variance of expenditure of customers at a supermarket, based on the data in file `supermarket.xls`.

- Compute a 95% confidence interval for the variance, assuming that the random variable *total expenditure* follows a normal distribution. (Keep in mind that we know it does not!)
- A resampling method (bootstrap) shows that the standard error of the point estimator for the variance is about 17. Use this information to compute an approximate 95% confidence interval for the variance, according to “point estimator  $\pm 2 \times$  its standard error”.
- Compare the two confidence intervals in (a) and (b). How do you explain their difference? Which one would you rather trust?