# Statistics for the Life Sciences

Math 20340 Section 01, Fall 2009

Homework 7 Solutions

- **8.1**: A $95\%$ margin of error for a point estimate of a parameter is a measure of the maximum amount by which the estimate will differ from the actual parameter $95\%$ *of the time in which the procedure that produces the point estimate is performed.*
- **8.3**: In each case, margin of error is $\pm 1.96 * \frac{\sigma}{\sqrt{30}}$.
  - **a**: $\pm .16...$
  - **b**: $\pm .339...$
  - **c**: $\pm .438...$
- **8.4**: Increasing the population variance increases the margin of error.
- **8.5**:
  - **a**: $\pm .554...$
  - **b**: $\pm .175...$
  - **c**: $\pm .055...$
- **8.6**: Increasing the sample size decreases the margin of error.
- **8.9**: In each case, margin of error is $\pm 1.96 * \frac{\sqrt{pq}}{\sqrt{100}}$.
  - **a**: $\pm .0588$
  - **b**: $\pm .0898...$
  - **c**: $\pm .098$
  - **d**: $\pm .0898...$
  - **e**: $\pm .0588$
  - **f**: $p = .5$ produces largest margin of error.
- **8.14**: Here $n = 75$ (large enough to assume a normal distribution for the sample mean), $\bar{x} = 4.2$ and $s = 1.5$, so the sample distribution standard error (SE) is $s/\sqrt{n} = .173....$ So an estimate of the average biomass per square meter is $4.2$ with (95%) margin of error $\pm 1.96 * .173 = \pm .339....$

- **8.17**:
  - **a**: Point estimate .51; (95%) margin of error $\pm 1.96 * \sqrt{(.51) * (.49)/900} = \pm .03266...$
  - **b**: To be able to apply the margin of error estimate to *all* questions, we shouldn't use $\hat{p}$ as an estimate for $p$, but instead use the "worst-case" estimate $p = .5$, leading to the largest possible margin of error (which, rounded down, is $\pm 3\%$).
- **8.19**:
  - **a**: No. For example, sample is skewed towards people with plenty of time (and money!) on their hands.
  - **b**: Valid results might be obtained (and so MOE might be interesting) if the population you are interested in is those people who tend to respond to 900 number opinion polls; but if you are interested in inferring results for the population as a whole, then the results of the survey are probably not valid (and therefore the MOE is meaningless).
- **8.24 b**: $\alpha = .1$ so $z_{\alpha/2} = z_{.05} = 1.645$. $90\%$ confidence interval is $1049 \pm 1.645 * \sqrt{51}/\sqrt{65} = 1049 \pm 1.457$.
- **8.25**: $\hat{p} = 263/300 = .876...$. $90\%$ confidence interval is $.876 \pm 1.645 * \sqrt{.876 * .124/300} = .876 \pm .031...$. Interpretation: $90\%$ of the time in which this procedure is used to produce a confidence interval, $p$ (the actual population proportion) will lie in the produced interval.
- **8.27**: $95\%$ confidence interval is $\bar{x} \pm 1.96 * (\sigma/\sqrt{n})$, so the width of the interval is $2 * 1.96 * (\sigma/\sqrt{n}) = 39.2/\sqrt{n}$.
  - **a**: $n = 100$, width is $3.92$
  - **b**: $n = 200$, width is $2.77...$
  - **c**: $n = 400$, width is $1.96$
- **8.28**:
  - **a**: Doubling the sample size decreases the width of the confidence interval by a factor of $\sqrt{2}$.
  - **b**: Quadrupling the sample size decreases the width of the confidence interval by a factor of 2.
- **8.32**:
  - **a**: $.54 \pm 1.96 * \sqrt{.54 * .44/400} = .54 \pm .047...$
  - **b**: $.3 \pm 1.96 * \sqrt{.3 * .7/350} = .3 \pm .048...$
- **8.36**:

- **a**: Perhaps not; people tend to be either very computer literate or very computer illiterate. So its possible that online transaction times are very low for some people (much less than $4.5$ minutes) and very high for others (much more than $4.5$ minutes). Also, the older generation (who tend to be the less computer savvy) perhaps tend also to have slower connections, whereas younger people (who tend to be more comfortable with computers) tend to have much faster connections. Both of these factors might lead to a distribution of times that has *two* peaks (one below $4.5$ minutes, one above $4.5$ minutes), rather than one, as a mound-shaped distribution would have.
- **b**: Central Limit Theorem says that whatever the population distribution, the distribution of the sample mean is approximately normal as long as the sample size is large enough. Here the sample is $50$, comfortably large enough.
- **c**: $95\%$ confidence interval: $4.5 \pm 1.96 * (2.7/\sqrt{50}) = 4.5 \pm .748$.