Ma 494 — Theoretical Statistics

Solutions for Problem Set #4 — Due March 3, 2010

Prof. Sawyer — Washington University

NOTE: 5 problems on 2 pages.

1. Note $P\big(\overline{X}_n \in (2.5, 3.5)\big) = P\big(|\overline{X}-3| < 0.50\big) \ge 0.99$ is equivalent to $P\big(|\overline{X}-3| > 0.50\big) \le 0.01$. Since X_i are $N(3,1)$, $\overline{X}$ is $N(3, 1/n)$, so that $\sqrt{n}\big(\overline{X}-3\big) \approx Z$ where Z is standard normal. Hence

$$\begin{aligned} P\big(|\overline{X}-3| > 0.50\big) &= P\big(\sqrt{n}|\overline{X}-3| > 0.50\sqrt{n}\big) \\ &= P\big(|Z| > 0.50\sqrt{n}\big) = 2P(Z > 0.50\sqrt{n}) \ \le \ 0.01 \end{aligned}$$

if and only if $P(Z > 0.50\sqrt{n}) \le 0.005$. By Table A1 pages 851–852, this is equivalent to $0.50\sqrt{n} \ge 2.575$ or $n \ge (2.575/0.50)^2 = 5.150^2 = 26.53$, or $n \ge 27$ since n must be an integer.

2. Here $X_1, \dots, X_n$ for $n = 7$ are independent $N(\mu, 1.2^2)$ with $\overline{X} = 0.80$. Thus $Z = \big(\overline{X}-\mu\big)/(1.2/\sqrt{7})$ is standard normal. Since $P(-1.960 < Z < 1.960) = 0.95$, it follows as in Example 5.3.1, that

$$\left(0.80 - 1.960\frac{1.2}{\sqrt{7}},\ 0.80 + 1.960\frac{1.2}{\sqrt{7}}\right) \ = \ (-0.0890, 1.6890)$$

is a symmetric two-sided 95% confidence interval for μ. Since

$$\begin{aligned} P(Z < 1.6445) &= 0.95 = P\big(\big(\overline{X}-\mu\big)/(1.2/\sqrt{7}) < 1.6445\big) \\ &= P\big(\big(\mu-\overline{X}\big)/(1.2/\sqrt{7}) > -1.6445\big) = P\left(\mu > \overline{X} - 1.6445\frac{1.2}{\sqrt{7}}\right) = 0.95 \end{aligned}$$

it follows similarly that $\big(0.80 - 1.6445(1.2/\sqrt{7}),\ \infty\big) = (0.0541,\ \infty)$ is a lower 95% one-sided confidence interval for μ. Thus, in these senses, with 95% confidence, we can say that $\mu > 0$, but not that $\mu \ne 0$.

3. Let p be the true proportion of unsatisfactory tuna salads at these establishments. Then $\widehat{p} = 179/220 = 0.8136$ where $\widehat{p}$ is approximately normal with distribution $N\big(p,\ p(1-p)/n\big)$. An approximate 90% (not 95%) symmetric confidence interval is

$$\begin{aligned} &\left(\widehat{p} - 1.6445\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}},\ \widehat{p} + 1.6445\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}\right) \\ &= \ (0.8136 - 0.0432,\ 0.8136 + 0.0432) \ = \ (0.7705, 0.8568) \end{aligned}$$

4. Let N be the number of votes that Tom Foley obtained before the $n = 14{,}000$ absentee votes are counted and Q be the number of votes that he obtains from the absentee ballots. Let p be the probability that an individual absentee votes for Foley. We need the smallest value of p so that Foley has a 20% chance of winning the election.

If Foley obtained N votes among the regular ballots, and the absentee ballots are split between Foley and Nethercut (that is, no blank or spoiled ballots and no-one voted for a third candidate, such as themselves), then the final number of votes for the two candidates is

Tom Foley	$N + Q$
George Nethercut	$N + 2174 + 14{,}000 - Q$

Foley will win if and only if $Q > 2174 + 14000 - Q$ or $2Q > 16174$ or $Q > 8087$. Since p will be reasonably close to 1/2, we can use the conservative approximation $Q/n \approx N\big(p, 1/(4n)\big)$. Thus $Q \approx N\big(np, n/4)$ and $Z = (Q - np)/\sqrt{n/4}$ is approximately standard normal. Hence we need

$$P(Q > 8087) \;=\; P\big(Q - np > 8087 - np\big) \;=\; P\Big(Z > (8087 - np)/\sqrt{n/4}\Big) = 0.20$$

Since $P(Z > 0.8416) = 0.20$, we solve $0.8416 = (8087 - 14000p)/\sqrt{14000/4}$ or

$$p \;=\; \Big(8087 - 0.8416 \times \sqrt{14000/4}\Big)/14000 \;=\; 0.5741$$

This is only slightly smaller than the actual fraction of absentee ballots $p_0 = 8087/14000 = 0.5776$ that Foley needs to win.

5. It is sufficient to find the smallest n such that $\big|(X/n) - p\big| > 0.05$ with probability 0.01 or smaller for any p. Here X is binomial $\mathrm{bin}(n, p)$ and we can use the conservative approximation $(X/n) \approx N\big(p, 1/(4n)\big)$, so that the distribution of $(X/n) - p \approx N\big(0, 1/(4n)\big)$ does not depend on p. Thus we need

$$\begin{aligned} P\big(\big|(X/n) - p\big| > 0.05\big) \;&=\; P\big(\sqrt{4n}\big|(X/n) - p\big| > 0.05\sqrt{4n}\big) \\ &=\; P\big(|Z| > 0.05\sqrt{4n}\big) \;=\; 2P\big(Z > 0.05\sqrt{4n}\big) \;<\; 0.01 \end{aligned}$$

Since $P(Z > 2.5758) = 0.005$, we need $0.05\sqrt{4n} > 2.5758$ or $n > (2.5758/0.10)^2 = (25.758)^2 = 663.47$. Thus we must have $n \geq 664$.