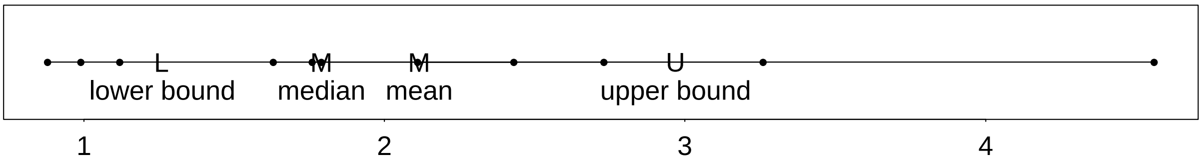
ST102 Class 15 – Solutions to Additional exercises

1. (a) We have $\bar{x} = 2.115$, and the top 0.5th percentile of $N(0,1)$ is $z_{0.005} = 2.576$. Hence a 99% confidence interval for μ is:

$$\bar{x} \pm z_{0.005} \times \frac{\sigma}{\sqrt{n}} = 2.115 \pm 2.576 \times \frac{1.1}{\sqrt{11}} \quad \Rightarrow \quad (1.260, 2.969).$$

   (b) The sample median is 1.79, and a plot of the data is shown below.



   (A hand-drawn graph is perfectly acceptable.)

   (c) Let the length of the interval be such that $2 \times z_{0.005} \times \sigma/\sqrt{n} \leq 0.4$. Hence:

$$n \geq \frac{2^2 \times (z_{0.005})^2 \times \sigma^2}{(0.4)^2} = \frac{4 \times (2.576)^2 \times (1.1)^2}{0.16} = 200.73.$$

   Hence the sample size should be at least as large as $n = 201$ in order to ensure that the width of the interval is not greater than 0.4, i.e. at least 190 *more* observations are required.

Note that the confidence intervals obtained here are very wide, leading to inaccurate estimates. This is due to the small sample size used for the estimation.

2. Since n is large, then approximately:

$$\frac{X}{n} \sim N\left(\pi, \frac{\pi(1-\pi)}{n}\right).$$

Since:

$$P(-0.67 < Z < 0.67) = 0.4971 < 0.50 = P(0 < Z < \infty)$$

where $Z \sim N(0,1)$, then:

$$\left(\frac{X}{n}, \infty\right).$$

has a slightly greater probability of containing π. However, the difference is marginal as both intervals have coverage probability of approximately 50%.

3. (a) In this case:

$$n \geq \frac{(z_{0.02})^2}{4 \times (0.05)^2} = \frac{(2.05)^2}{4 \times (0.05)^2} = 420.25$$

   hence $n = 421$.

(b) In this case:

$$n \geq \frac{(z_{0.04})^2}{4 \times (0.04)^2} = \frac{(1.75)^2}{4 \times (0.04)^2} = 478.52$$

hence $n = 479$.

Therefore, case (b) demands the larger sample size.

4. Since $z_{0.005} = 2.576$, we have *approximate* (by the central limit theorem) confidence interval endpoints of:

$$2.576 \times \sqrt{\frac{p(1-p)}{n}}.$$

which is (noting that the maximum value of $p(1-p)$ is when $p = 0.50$):

$$\leq 2.576 \times \sqrt{\frac{1}{4n}} \leq 0.005$$

so we take:

$$n \geq \frac{(2.576)^2}{4 \times (0.005)^2} = 66{,}357.76$$

which in practice is 66,358 – quite large indeed!

5.* As defined, R is a random variable, and $R \sim \text{Bin}(n, \pi)$, so that $\text{E}(R) = n\pi$ and hence $\text{E}(R/n) = \pi$. It also follows that:

$$1 - \text{E}\left(\frac{R}{n}\right) = \text{E}\left(1 - \frac{R}{n}\right) = \text{E}\left(\frac{n-R}{n}\right) = 1 - \pi.$$

So the first obvious guess is that we should try $R/n \times (1 - R/n) = R/n - (R/n)^2$. Now:

$$n\pi(1-\pi) = \text{Var}(R) = \text{E}(R^2) - (\text{E}(R))^2 = \text{E}(R^2) - (n\pi)^2.$$

So:

$$\text{E}\left(\left(\frac{R}{n}\right)^2\right) = \frac{1}{n^2}\,\text{E}(R^2) = \frac{1}{n^2}\left(n\pi(1-\pi) + n^2\pi^2\right)$$

$$\Rightarrow \quad \text{E}\left(\frac{R}{n} - \left(\frac{R}{n}\right)^2\right) = \frac{1}{n}\,\text{E}(R) - \frac{1}{n^2}\left(n\pi(1-\pi) + n^2\pi^2\right)$$

$$= \frac{n\pi}{n} - \frac{n^2\pi^2}{n^2} - \frac{\pi(1-\pi)}{n}$$

$$= \pi - \pi^2 - \frac{\pi(1-\pi)}{n}.$$

However, $\pi(1-\pi) = \pi - \pi^2$, so:

$$\text{E}\left(\frac{R}{n} - \left(\frac{R}{n}\right)^2\right) = \pi(1-\pi) - \frac{\pi(1-\pi)}{n} = \pi(1-\pi) \times \frac{n-1}{n}.$$

It follows that:

$$\pi(1-\pi) = \frac{n}{n-1} \times \mathrm{E}\left(\frac{R}{n} - \left(\frac{R}{n}\right)^2\right) = \mathrm{E}\left(\frac{R}{n-1} - \frac{R^2}{n(n-1)}\right).$$

So we have found the unbiased estimator of $\pi(1-\pi)$, but it could do with tidying up! When this is done, we see that:

$$\frac{R(n-R)}{n(n-1)}$$

is the required unbiased estimator of $\pi(1-\pi)$.