Math Camp — Day 4 Exercises

Worked Solutions

Exercise 1 — A weighted estimator of μ

We have: $\tilde{Y} = \frac{1}{2}Y_1 + \frac{1}{3}Y_2 + \frac{1}{6}Y_3$.

NB: as Y_1, Y_2, Y_3 is a random sample, we can write $Y_1, Y_2, Y_3 \overset{iid}{\sim} N(\mu, \sigma^2)$.

a) Unbiasedness.

$$E(\tilde{Y}) = \tfrac{1}{2}\mu + \tfrac{1}{3}\mu + \tfrac{1}{6}\mu = \mu\left(\tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{6}\right) = \mu \cdot \frac{3+2+1}{6} = \mu.$$

Since $E(\tilde{Y}) = \mu$, the estimator is unbiased. (Note the weights sum to 1.)

b) Variance vs. the sample mean. By independence,

$$\mathrm{Var}(\tilde{Y}) = \left(\tfrac{1}{2}\right)^2\sigma^2 + \left(\tfrac{1}{3}\right)^2\sigma^2 + \left(\tfrac{1}{6}\right)^2\sigma^2 = \sigma^2\left(\tfrac{1}{4} + \tfrac{1}{9} + \tfrac{1}{36}\right) = \sigma^2 \cdot \frac{9+4+1}{36} = \frac{14}{36}\sigma^2 = \frac{7}{18}\sigma^2 \approx 0.389\,\sigma^2.$$

The sample mean $\bar{Y} = \frac{1}{3}(Y_1 + Y_2 + Y_3)$ has

$$\mathrm{Var}(\bar{Y}) = \frac{\sigma^2}{3} = \frac{6}{18}\sigma^2 \approx 0.333\,\sigma^2.$$

Thus $\mathrm{Var}(\tilde{Y}) > \mathrm{Var}(\bar{Y})$.

c) Comparison. Both estimators are unbiased, but $\bar{Y}$ has the smaller variance, so $\bar{Y}$ is more efficient. Hence $\tilde{Y}$ is **not** as good an estimator as $\bar{Y}$. (Equal weights $1/n$ minimise the variance among unbiased linear combinations.)

Exercise 2 — Survey of 400 voters (215 in favour)

a) Point estimate.

$$\hat{p} = \frac{215}{400} = 0.5375.$$

b) Standard error.

$$\mathrm{SE}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.5375 \cdot 0.4625}{400}} = \sqrt{0.0006215} \approx 0.0249.$$

NB: the standard error of the mean corresponds to the standard deviation $\hat{\sigma}$ divided by $\sqrt{n}$.

c) Two-sided test $H_0 : p = 0.5$ vs. $H_1 : p \neq 0.5$.

$$t = \frac{\hat{p} - 0.5}{\mathrm{SE}(\hat{p})} = \frac{0.0375}{0.0249} \approx 1.50.$$

Using the table, $\Phi(1.50) = 0.9332$, so the two-sided p-value is

$$p\text{-value} = 2\big[1 - \Phi(1.50)\big] = 2(0.0668) = 0.1336 \approx 0.13.$$

d) One-sided test $H_0 : p = 0.5$ vs. $H_1 : p > 0.5$. Same statistic $t \approx 1.50$, but only the right tail counts:

$$p\text{-value} = 1 - \Phi(1.50) = 0.0668 \approx 0.067.$$

e) Why they differ. Under the two-sided alternative, evidence *either* above *or* below 0.5 counts, so both tails are included; the one-sided alternative uses only the upper tail. The two-sided p-value is therefore exactly twice the one-sided one.

f) Statistically significant evidence in favour? Even the (smaller) one-sided p-value, 0.067, exceeds 0.05, and the two-sided value is 0.13. In both cases $p\text{-value} > 0.05$, so we **fail to reject** H_0: the survey did *not* provide statistically significant evidence at the 5% level that voters favoured a unified health insurance.

g) 95% confidence interval ($z^{\text{crit}} = 1.96$).

$$\hat{p} \pm 1.96\ \mathrm{SE}(\hat{p}) = 0.5375 \pm 1.96(0.0249) = 0.5375 \pm 0.0488 = (0.489,\ 0.586).$$

h) 99% interval — wider or narrower? Wider: a higher confidence level uses a larger critical value ($z^{\text{crit}} = 2.576$ instead of 1.96), so the margin of error grows.

i) 5% test with no extra calculation. By the duality between confidence intervals and two-sided tests: the value 0.5 lies *inside* the 95% interval $(0.489, 0.586)$, so at the 5% level we **do not reject** $H_0 : p = 0.5$ — consistent with part (c)/(f).