Exam 2 Soultions — Fall 2010 — ACMS 20340

1. X is normal, so standardize and do a table look-up:

$$P(X < 0.5) = P(Z < \frac{0.5 - 0}{2} = P(Z < 0.25) = 0.5987$$

For your reference the answer choices in the test *should* have been (a) 0.6915, (b) 0.5987, (c) 0, (d) 0.8413, (e) 0.5

2. The probability of rolling a 6 is $p = 1/6$ and there are $n = 150$ trials. X is counting the number of 6's. This makes X is binomial with $\mu_X = np = 25$ and $\sigma_X = \sqrt{npq} = 4.564$. We need to approximate X with a normal rv. The approximation formula is $P(X \geq a) \approx P(Z \geq (a - 0.5 - \mu_X)/\sigma_X)$. The probability we need has a strict inequality, so first take care of that, and then approximate:

$$P(X > 30) = P(X \geq 31) \approx P(Z \geq \frac{30.5 - 25}{4.56})$$

3.1 It is clear $f(x) \geq 0$ for every x. We just need the area under f to be equal to 1.

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_0^c \tfrac{2}{9} x\, dx = \tfrac{1}{9} c^2$$

Solving $\frac{1}{9}c^2 = 1$ gives $c = \sqrt{9} = 3$.

3.2 Using $c = 3$, the mean of X is then calculated as

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx = \int_0^3 \tfrac{2}{9} x^2 = \frac{2}{27}\left[3^3 - 0\right] = 2$$

3.3 First calculate the variance. Here I use the formula $Var(X) = E(X^2) - \mu^2$. So

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\, dx = \int_0^3 \tfrac{2}{9} x^3 = \frac{2}{9 \cdot 4}\left[3^4 - 0^4\right] = \frac{9}{2}$$

And then $Var(X) = \frac{9}{2} - 2^2 = \frac{1}{2}$. Taking the square root: $\sigma_X = \sqrt{Var(X)} = 1/\sqrt{2}$.

4.1 The question is to compare the percentage of young voters with the overall percentage of voters. So the appropriate statistic is $\hat{p}_1 - \hat{p}_2$.

4.2 We want our estimate to be within 3% of the true value 90% of the time. This is the same as saying we want our 90% confidence interval to have a radius of 0.03. The 90% confidence interval has $\alpha = 0.1$ and $z_{\alpha/2} = z_{0.05} = 1.645$, and the interval is calculated as $(\hat{p}_1 - \hat{p}_2) \pm z_{0.05} SE(\hat{p}_1 - \hat{p}_2)$. Our goal is to find n so that

$$z_{0.05} SE(\hat{p}_1 - \hat{p}_2) = 0.03$$

We are assuming the two samples have the same size n. We also know the sample error is maximized when both $p = 0.5$ and $q = 0.5$. Since we want to find a value of n to satisfy the above equation, lets assume the worst. So then

$$\begin{aligned} z_{0.05} SE(\hat{p}_1 - \hat{p}_2) &= (1.645)\sqrt{\frac{p_1 q_1}{n} + \frac{p_2 q_2}{n}} \\ &\leq (1.645)\sqrt{\frac{0.5^2 + 0.5^2}{n}} \\ &= \frac{(1.645)\sqrt{2/n}}{2} \end{aligned}$$

Set the above equal to 0.03 and solve for n:

$$(1.645)\sqrt{2}/2 = 0.03\sqrt{n}$$

So $\sqrt{n} = 38.77$ and $n = 1503.3$, which rounds up to 1504.

4.3 Calculate the two proportions as $\hat{p}_1 = 83/257 = 0.323$ and $\hat{p}_2 = 94/313 = 0.300$, giving a difference of 0.023. The standard error is found to be

$$\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} = \sqrt{\frac{(0.323)(0.677)}{257} + \frac{(0.3)(0.7)}{313}} = 0.039$$

For a 90% confidence interval we use $z_{0.05} = 1.645$, and so the interval is $0.023 \pm (1.645)(0.039)$.

5 This is a straightforward confidence interval problem. The problem is asking for an interval estimate for the average weight of the boxes, i.e. an interval of possible values of μ. From the problem statement $n = 50, \bar{x} = 4.89, s = 0.3$. Calculate $SE(\bar{x}) = s/\sqrt{n} = 0.424$. The 95% confidence interval means $\alpha = 0.05$, so $z_{\alpha/2} = z_{0.025} = 1.96$. The interval is $\bar{x} \pm z_{0.025} SE(\bar{x})$. Some arithmetic gives $4.807 < \mu < 4.973$.

For the second part, we desire the boxes to have an average weight of 5 pounds. The confidence interval constructed in the first part gives an estimate of the true average box weight. Since 5 pounds is not in the interval, we conclude the machine is probably not working correctly. We say it is *probably* not working correctly since it is possible we had an anomalous sample for which true average weight is not in this interval. But, since we constructed the interval at the 95% confidence level, the probability of having a bad sample is only 5%. Further investigation into the machine is warranted.

In particular, the saying the interval is at the 95% level just means if we were to take many samples and construct similar confidence intervals for each, we would expect the true average μ to be in about 95% of these intervals.

6 Let X be a uniformly distributed random variable on the interval $[0, 10]$. This means X is a continuous rv with the density function

$$f(x) = \begin{cases} \frac{1}{10}, & \text{if } 0 \le x \le 10 \\ 0 & \text{otherwise} \end{cases}$$

We let $\{x_1, \ldots, x_{60}\}$ be our waiting times for 60 (one-way) trips. And then $T = x_1 + \cdots + x_{60}$ is the total time spent waiting. Since T is the sum of 60 samples, the central limit theorem says T is approximately normal with mean $\mu_T = n\mu_X = 60\mu_X$, and standard deviation $\sigma_T = \sigma_X \sqrt{n} = \sigma_X \sqrt{60}$.

We first need to find μ_X; an integral does this:

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x f(x)\, x = \int_0^{10} \frac{x}{10}\, dx = \frac{100}{20} = 5$$

And so $\mu_T = 60\mu_X = 300$.

I calculate the standard deviation of X by first finding

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) = \int_0^{10} \frac{x^2}{10}\, dx = \frac{100}{3}$$

And then $Var(X) = E(X^2) - \mu_X^2 = 100/3 - 25 = 8.3$, so $\sigma_X = \sqrt{Var(X)} = \sqrt{8.3} = 2.88$. Now that we have the standard deviation of X, the standard deviation of T is $\sigma_T = \sigma_X \sqrt{60} = 22.32$.