STATISTICS FOR ENGINEERS
ASSIGNMENT X ANSWERS
NOVEMBER 5, 2010

Problem 10.1

Let $X \sim \mathcal{N}(\mu = 10, \sigma^2 = 10)$. What is the probability that X is larger than 13?

Solution: We are given that $X \sim \mathcal{N}(\mu = 10, \sigma^2 = 10)$. Thus, in order to use the tables in the book, we transform X into a Z variable: $Z = \frac{X-\mu}{\sigma^2}$. This z is now distributed $Z \sim \mathcal{N}(0, 1)$. From the formula, $z = \frac{13-10}{10} = 0.3$. Going to our tables in the book, we find $\mathbb{P}[Z \geq 0.3] = 0.6179$. Thus, the probability that X is larger than 13 is approximately 0.6179. $\diamond$

- The z-score formula is $z = \frac{X-\mu}{\sigma}$.
- As such, $\mathbb{P}[X > 13] = 1 - \mathbb{P}[X \leq 13] = 1 - 0.8289 = 0.1711$

Problem 10.2

In a sample of 43 power companies in the southern United States, 20 indicated that state rules forced them to change their plans for building coal-fired power plants. In a sample of 93 power companies in the western United States, 55 indicated that state rules forced them to change their plans for building coal-fired plants. Can we conclude that the proportion of power companies in the South that had to change their plans for building coal-fired plants is less than those in the West?

Solution: This is a difference of proportions test with independent samples. Thus, we use the formula on Page 428 to compute a z-score:

$$\begin{aligned} z &= \frac{\hat{p}_X - \hat{p}_Y}{\sqrt{\hat{p}(1-\hat{p})(1/n_X + 1/n_Y)}} \\ &= \frac{0.4651 - 0.5914}{\sqrt{0.5662(1-0.5662)(1/43+1/93)}} \\ &= 0.87 \end{aligned}$$

Now, going to our z-table, we find the probability $\mathbb{P}\left[Z \geq 0.87\right] = 0.1922$. Thus, as this number is greater than α, we can conclude that the two proportions are the same. $\diamond$

- The null hypothesis is $H_0 : p_W \geq p_S$. The null hypothesis ALWAYS contains the 'equals' portion.
- The $\hat{p}$ is calculated incorrectly. $\hat{p} = 0.5515$.
- With this, $z = 1.377$. The one-tailed probability is $p = 0.0838$. Thus, we cannot reject.
- Since the samples have the southern proportion less than the western proportion, we know $p < 0.500$.

Problem 10.3

Let us (somehow) know that the reaction times of a mouse have a symmetric distribution. A psychology student ran an experiment where she measured the time between when she screamed at the mouse and when it jumped. She got the following reaction times for Mouse 1: 36, 28, 29, 20, 38. She got the following for Mouse 2: 34, 41, 35, 47, 49, 46. Can we conclude that there is a significant difference in the reaction times between the two mice?

Solution: We are given that the distributions are symmetric, although they are probably not Normal, since we have some outliers. Thus, we will use a non-parametric test, the Mann-Whitney test.

Ranking the reaction time gives us 20, 28, 29, 34, 35,36, 38, 41, 46, 47, and 49. These, respectively, come from Mouse number 1, 1, 1, 2, 2, 1, 1, 2, 2, 2, and 2. The sum of the ranks for Mouse 1 is 19.

The sum of the ranks for Mouse 2 is 47. The difference between the two is $47 - 19 = 26$. Thus, $W = 26$. Using Table A.5 in the back of the book, with $n = 6$ and $m = 5$, we get that the p-value is greater than 0.0628. Thus, at the usual level ($\alpha = 0.05$), we can conclude that the mice have the same reaction times. ◇

- We discover that $47 - 19 = 28$, not 28.
- The Mann-Whitney test does not have you subtract the two values. The W statistic is one or the other.
- This is a two-tailed test. As such, we need to know what number Table A.5 actually gives us area under one tail (like the other tables).
- Thus, you will need to double the table probability (value you read off in the table).

Problem 10.4

A scale is calibrated by weighing a 1.000kg test weight 60 times. Those readings are Normally distributed and have a mean of 1.006kg and a standard deviation of 2g. Find the p-value for testing $H_0 : \mu = 1.000kg$.

Solution: We must first calculate the z-value for this problem. Here, $z = \frac{0.6}{2} = 0.3$. Looking in our z-tables, we find $p = 0.6179$. Thus, since this is greater than 0.05, we accept the null hypothesis that the scale is perfectly calibrated. $\diamond$

- The units for the measurements are in kg. The units for the standard deviation are in g. Thus, either do the test with $\bar{x} = 1006$g or $s = 0.002$kg. Be consistent.
- The correct z-score formula needs an n value in it, since we are looking at sample *means*: $z = \frac{\bar{x}-\mu}{s/\sqrt{n}}$.
- We do not accept the null hypothesis. It is obvious that the scale is *not* perfectly calibrated. We can, in fact, conclude that the nll hypothesis is extremely unlikely ($z \approx 23$; $p \approx 0$).

Problem 10.5

Election returns from the 2009 presidential election in Iran have the following first digit frequencies:

⋮

The Center for Electoral Forensics performed the Benford Test on the frequencies and concluded that there was a statistically large departure from expectation for these digit frequencies. Were they correct?

Solution: The null hypothesis is that the digit frequencies follow the Benford distribution, which is provided in the final problem of Examination One. To test this hypothesis, we will perform a chi-squared test. The expected frequencies and relevant calculations are in the following table.

Digit	1	2	3	4	5	6	7	8	9
Observed	66	60	53	46	40	33	25	16	11
Expected	40	40	40	40	40	40	40	40	40
$O - E$	26	20	13	6	0	-7	-15	-24	-29
$(O - E)^2$	676	400	169	36	0	49	225	576	841
$(O - E)^2/E$	16.9	10	4.225	0.9	0	1.225	5.625	14.4	21.025

The chi-square statistic is the sum of this last row, $X^2 = 74.3$. The degrees of freedom is $\nu = 8$. Therefore, the p-value is $6.8E-13$. As this is less than our usual $\alpha = 0.05$, we can reject the null hypothesis and agree with the conclusions of the Center for Electoral Forensics. ◇

- The expected values are incorrect. These are from the uniform distribution, not the Benford distribution.
- We, too, can agree with the conclusion of CEF. Hopefully, you notice that one can get the same conclusion while doing it wrong.