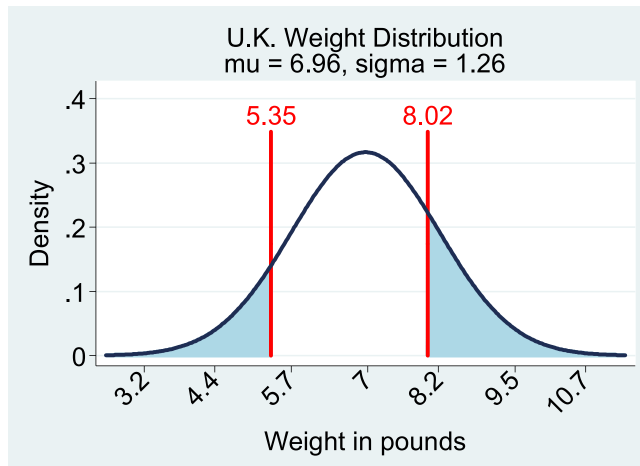
ECO220Y, Term Test #2: SOLUTIONS

January 13, 2017, 9:10 – 11:00 am

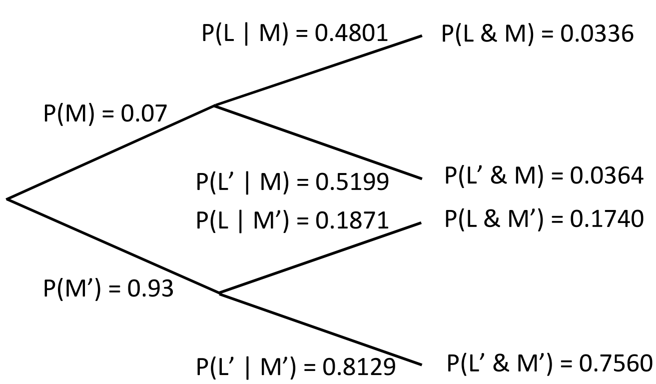
(1) (a) The mean birth weight in Finland is about 3,600 grams with a standard deviation of about 500 grams. We obtain these estimates by noting that the shape of the distribution is Normal. Because the Normal distribution is symmetric about the mean we find the centre value of the distribution as about 3,600. We use the Empirical Rule to find the approximate s.d.: it looks like about 99.7% of the births are between 2,100 and 5,100, which would imply a s.d. of about 500 because the Empirical Rule states that about 99.7% of the observations should lie within plus and minus 3 standard deviations from the mean (i.e. 3,000/6 = 500). [Note: It is also possible to answer using the other two parts of the Empirical Rule: that about 68.3% lie within plus/minus one s.d. of the mean or that about 95.4% lie within plus/minus two s.d.'s of the mean.]

(b) P(X < 5.35) = 0.10 and P(X < 8.02) = 0.80 where X is the weight in pounds: $X \sim N(\mu, \sigma^2)$. From the Normal table obtain P(Z < -1.28) = 0.10 and P(Z < 0.84) = 0.80. Hence, $\frac{5.35-\mu}{\sigma} = -1.28$ and $\frac{8.02-\mu}{\sigma} = 0.84$. Solving two equations for two unknows we obtain $\mu = 6.96$ pounds and $\sigma = 1.26$ pounds.



(2) (a) 0.15 is a conditional probability: the probability that an advisor has a misconduct record conditional on being at some of the largest financial advisory firms in the U.S. The fact that P(Misconduct) = 0.07 is not equal to P(Misconduct | Largest firms) = 0.15, means that the size of the firm and misconduct are *not independent*. In other words, the chance that an advisor engages is misconduct *is* related to the size of the firm, with larger firms having a higher incidence of misconduct than smaller firms.

(b) This is a reversing the conditioning problem. We must obtain the relevant values from the abstract and Table 8a. Define M as the event of misconduct and define L as the event of leaving a firm. We must find P(M | L) = ?.



P(M | L) = P(M & L)/P(L)

P(L) = P(M & L) + P(M' & L) = 0.0336 + 0.1740 = 0.2076

P(M | L) = P(M & L)/P(L) = 0.0336/0.2076 = 0.162

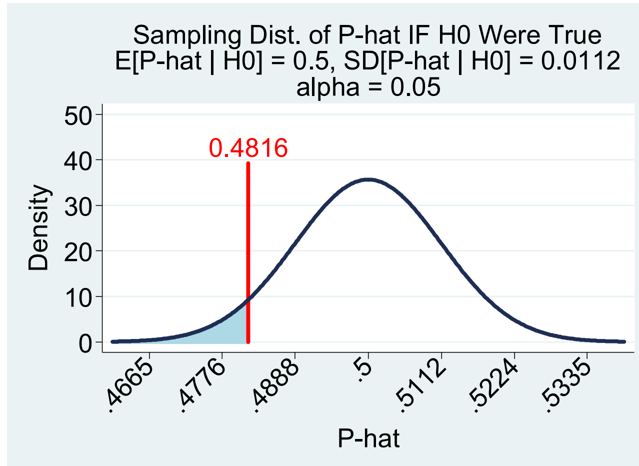
Hence, we should *not* be very suspicious of a financial advisor who leaves a firm: there is only a 16.2% chance that that advisor has committed misconduct.

(3)

$H_0: p = 0.50$

$H_1: p < 0.50$

Choose a significance level. One popular choice is $\alpha = 0.05$, which should convince a reasonable person. (Other acceptable choices would include $\alpha = 0.01$ and $\alpha = 0.10$.)



$P(Z < -1.645) = 0.05$

$$Z = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

$$-1.645 = \frac{\hat{P} - 0.5}{\sqrt{\frac{0.5(1-0.5)}{2{,}000}}}$$

$$critical\ value\ (c.v.) = \hat{P} = 0.5 - 1.645 * \sqrt{\frac{0.5(1-0.5)}{2{,}000}}$$

$$= 0.5 - 1.645 * 0.01118 = 0.5 - 0.0184$$
$$= 0.4816$$

Hence, we would need to see less than 48.2% of the random sample of 2,000 Torontonians in favor of the toll to convince a reasonable person that among all Torontonians less than half favor a toll on the DVP, which presumes a standard significance level, i.e. burden of proof, of 5%.

(4) (a) Graph #2 has a smaller variance than Graph #1. Graph #4 has a smaller variance than Graph #3. Graphs #1 and #2 show the sampling distribution of the sample mean for a sample size of n = 10 and n = 1,000, respectively. The larger the sample size the smaller the sampling error: $SD[\bar{X}] = \frac{\sigma}{\sqrt{n}}$. For Graphs #3 and #4, which show the sampling distribution of the sample median for a sample size of n = 10 and n = 1,000, respectively, we also have the concept that higher sample sizes mean less sampling error, but we do not have a formula like for the sample mean.

(b) The STATA summary shows the simulated sampling distribution of the sample median for a sample size of 10 ON public sector employees: how the sample median varies from one sample to another because of sampling error. The 5th percentile of the sampling distribution is 106.586, which means that there is a 5% chance that the sample median could be as small as $106,586 in a random sample of 10 ON public sector employees. Overall there is substantial sampling error in the sample median: it can be fairly different from the population median of $115,301 because there is considerable sampling error in a sample size of only 10 employees.

(5) (a) The first row reports the overall callback rate for all resumes. The rows after that break the results down by: (1) city, (2) occupation and degree requirements, (3) race and gender, and (4) average salary. This allows us to answer questions like: does the callback rate differ across cites? By occupation and degree requirements?, etc. Clearly there are cases where the callback rates vary substantially and others were they do not vary that much.

(b) A yes/no question about whether or not there is a difference requires a hypothesis testing approach to statistical inference: there is no direction specified so we must do a two-tailed test.

$H_0: p_{wh} - p_{nw} = 0$

$H_1: p_{wh} - p_{nw} \neq 0$

$$\hat{P}_{wh} = \frac{0.092 * 2{,}620 + 0.066 * 2{,}456}{2{,}620 + 2{,}456} = \frac{403}{5{,}076} = 0.0794$$

$$\hat{P}_{nw} = \frac{0.090 * 2{,}680 + 0.077 * 2{,}728}{2{,}680 + 2{,}728} = \frac{451}{5{,}408} = 0.0834$$

$$\bar{P} = \frac{X_{wh}+X_{nw}}{n_{wh}+n_{nw}} = \frac{403+451}{5{,}076+5{,}408} = 0.08146$$

$$z = \frac{\hat{P}_{wh} - \hat{P}_{nw}}{\sqrt{\frac{\bar{P}(1-\bar{P})}{n_{wh}} + \frac{\bar{P}(1-\bar{P})}{n_{nw}}}} = \frac{0.0794 - 0.0834}{\sqrt{\frac{0.08146(1-0.08146)}{5{,}076} + \frac{0.08146(1-0.08146)}{5{,}408}}} = \frac{-0.004}{0.005346} = -0.75$$

$$P - value = P(Z < -0.75) + P(Z > 0.75) = 2 * (0.5 - 0.2734) = 0.4532$$

These results are neither statistically significant nor economically significant. There is a tiny difference in the percent called back: 0.4 percentage points. Further, the P-value is very large, which means this type of tiny difference could easily be caused by sampling error even if there were absolutely no difference in the population proportions: it is not statistically significant at any reasonable significance level. (If we had found a difference we *would* have been able to infer causality because race and gender were randomly assigned.)

(c) A question about the size of the difference requires a confidence interval approach to statistical inference. We are not told which confidence level to use: a reasonable choice is to simply use the most common 95% confidence level. (Alternative reasonable choices would be a 99% or a 90% confidence level.)

$$(\hat{P}_{ge65} - \hat{P}_{lt35}) \pm z_{\alpha/2}\sqrt{\frac{\hat{P}_{ge65}(1-\hat{P}_{ge65})}{n_{ge65}} + \frac{\hat{P}_{lt35}(1-\hat{P}_{lt35})}{n_{lt35}}}$$

$$(0.048 - 0.105) \pm 1.96\sqrt{\frac{0.048(1-0.048)}{1{,}448} + \frac{0.105(1-0.105)}{2{,}497}}$$

$$-0.057 \pm 1.96 * 0.00832$$

$$-0.057 \pm 0.0163$$

$$LCL = -0.0733$$

$$UCL = -0.0407$$

The point estimate is that the callback rate is 5.7 percentage points lower for the most well-paid job postings ($65K or more) compared to the worst-paid postings (less than $35K): a difference of 10.5% being called back versus only 4.8%. Further, the margin of error is a fairly modest 1.6 percentage points. Overall we are 95% confident that callback rates among all job applicants are between 4.1 and 7.3 percentage points lower for the most well-paid job postings compared to the worst-paid postings. The callback rates a much higher (more than twice as high) for the worst-paid postings.