ECO220Y1Y, Test #2, Prof. Murdock SOLUTIONS

(1) (a)

$P(3{,}000 < W < 3{,}400) = ?$ To compute, use the Normal table.

$P(3{,}000 < W < 3{,}400) = P\left(0 < Z < \frac{3{,}400-3{,}148}{435}\right) + P\left(\frac{3{,}000-3{,}148}{435} < Z < 0\right)$

$\approx P(0 < Z < 0.58) + P(-0.34 < Z < 0) = 0.2190 + 0.1331 = 0.3521$

(b) $P(W < 3{,}975) = ?$

$P\left(Z < \frac{3{,}975-3{,}148}{435}\right) = P(Z < 1.90) = 0.5 + 0.4713 = 0.9713$, which means the 97th percentile

(c) $P(W < ?) = 0.14$ To compute, use the Normal table.

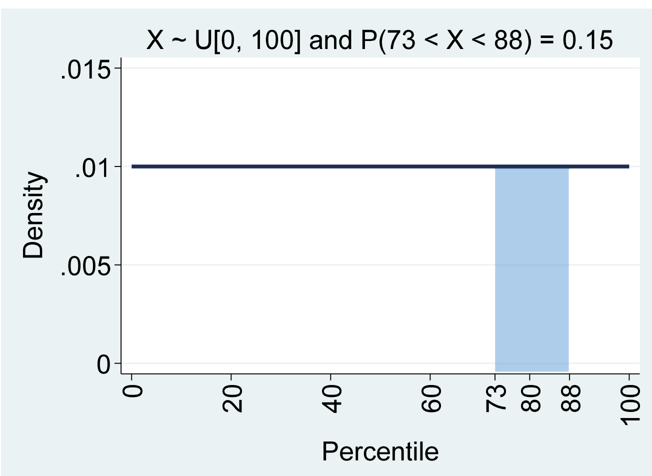
$P(Z < -1.08) = 0.14$

$Z = \frac{X-\mu}{\sigma}$

$-1.08 = \frac{X-3{,}148}{435}$

$X = 2678.2$ grams

(2) (a) $0.88 - 0.73 = 0.15$



(b) Yes. The population median (a parameter) is a salary of \$111,293, from Summary #1, and the sample median (a statistic) is only \$108,949. However, Summary #3 shows that for a sample size of 40 the sample median could easily be as small as \$108,949 because of sampling error: there is over a 25% chance of such a small sample median. This means that sampling error certainly is a plausible explanation.

(c) Scenario **B** would most impact the current value of the 5th percentile. Summary #3 shows the sampling distribution of the sample median, and a bigger sample size would mean that it is less spread out, making the 5th percentile bigger. The change in **A** is not expected to have much effect at all: changing the sample size does not change the shape or the spread of a sample (it will reflect the population aside from any random sampling error), and the sample 5th percentile (101.6699, which is a statistic) is already close to the population 5th percentile (101.2492, which is a parameter). The change in **C** is also not expected to have an effect: changing the number of simulation draws does not change the shape or the spread, it just gives a clearer image of the sampling distribution.

(3) (a) $\frac{0.075}{0.196} = 0.383$ [which means 38.3%]

(b) $P(\hat{P} > 0.45 \mid n = 500, p = 0.382) = ?$

$$P\left(Z > \frac{0.45-0.382}{\sqrt{\frac{0.382(1-0.382)}{500}}}\right) = P\left(Z > \frac{0.068}{0.0217}\right) = P(Z > 3.13) = 0.5 - 0.4991 = 0.0009$$

No, it is not plausible that over 45% of the sample would be 55 years and up, when only 38.2% of the population is: this is quite a large sample size (500) and the probability of being off by that much – nearly 7 percentage points off – is a bit less than 1%, meaning it is not a particularly plausible explanation.

[Note: Some students may set it up in terms of X, instead of $\hat{P}$. That is ok and yields the exact same answer:
$P(X > 225) = P\left(Z > \frac{225-191}{\sqrt{500*0.382*(1-0.382)}}\right) = P\left(Z > \frac{34}{10.865}\right) = P(Z > 3.13) = 0.5 - 0.4991 = 0.0009$]

[Note: Some students may do the continuity correction and find $P(X > 225.5)$ or $P(\hat{P} > 0.451)$. That is ok and yields a probability of about 0.0008, which is very close to 0.0009.]

(c) $P(X < 4 \mid n = 5, p = 0.894) = ?$

$P(X < 4) = 1 - P(X = 5) - P(X = 4)$

Use $p(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$

$p(5) = \frac{5!}{5!(5-5)!} 0.894^5 (1 - 0.894)^{5-5} = 0.894^5 = 0.571$

$p(4) = \frac{5!}{4!(5-4)!} 0.894^4 (1 - 0.894)^{5-4} = 5 * 0.894^4 (1 - 0.894) = 0.339$

$P(X < 4) = 1 - 0.571 - 0.339 = 0.09$ Hence, it seems somewhat plausible that this could happen by chance: there is 9% chance of such a small number being employed – 60% or less – even though this group is highly likely to be employed (89.4%).

[Note: Some students may set it up in terms of $\hat{P}$, instead of X. That is ok and yields the exact same answer.]

(4) (a) …, the rate of admissions for typical African American applicants is 2.7 percentage points higher and 55% higher than for typical white applicants. For both races this focuses on students who are not recruited athletes, legacies, on the Dean's interest list, or children of faculty or staff at Harvard: students who *are* in those special groups have a large advantage in admissions. Hence, *overall*, we cannot say that African American applicants have a large advantage over white applicants.

(b) To compute the standard deviation for each, plug into $SD[\hat{P}] = \sqrt{\frac{p(1-p)}{n}}$ and then multiple by 100.

For a random sample of 10 white legacy applicants, the expected percent admitted is 37 with a standard deviation of 15. For a random sample of 20 admitted students, the expected percent that are recruited athletes is 10 with a standard deviation of 7. For a random sample of 30 white admitted students, the expected percent that are non-ALDC students is 57 with a standard deviation of 9.

(5) (a)

- $P(DQ2 \mid NQ4) = \frac{P(DQ2 \,\&\, NQ4)}{P(NQ4)} = \frac{0.032}{0.20} = 0.16$
- $P(NQ4 \mid DQ2) = \frac{P(DQ2 \,\&\, NQ4)}{P(DQ2)} = \frac{0.032}{0.20} = 0.16$
- No, in general the order matters with conditional probabilities. The special thing about Table 3 is that it divides teacher-years into five quintiles: by definition, the probability of being in any quintile is one-fifth (20%). Hence, either way, we rescale the same joint probability by 0.20 to get the conditional probability.

(b) All cells would say 4% (=100*0.2*0.2): every one of the 25 combinations would be equally likely if there were no relationship at all between the value-added measures.

(c) …, for teachers each year of English language arts, 1.4% of the teacher-years are both in the fifth (top) quintile for the value-added measure for students with disabilities *and* in the first (bottom) quintile for the value-added measure for students without disabilities. This is one of the least likely of the 25 possible combinations: of all teacher-years, not many are especially good for students with disabilities and especially poor for students without disabilities.

(d) For example, disjoint events are the event that a teacher-year is in the first value-added quintile in math for students with disabilities and the event that a teacher-year is in the second value-added quintile in math for students with disabilities.

[There are four types of examples that work for mutually exclusive (disjoint) events: (1) being in two different quintiles for the non-SWD measures of value-added for math, (2) being in two different quintiles for the SWD measures of value-added for math, (3) being in two different quintiles for the non-SWD measures of value-added for English language arts, (4) being in two different quintiles for the SWD measures of value-added for English language arts. You CANNOT mix-and-match because no other type of combination is mutually exclusive.]

These events are NOT independent: if a teacher-year is in the first quintile then we know for sure they are not in the second quintile and independence requires P(Q2 | Q1) = P(Q2) and that is violated because zero is not equal to 0.2.