ECO220Y1Y, Test #2, Prof. Murdock SOLUTIONS

December 3, 2021, 9:10 – 11:00 am

(1) $E[X] = 0.33 * 0 + 0.49 * 1 + 0.18 * 2 = 0.85$

$V[X] = 0.33 * (0 - 0.85)^2 + 0.49 * (1 - 0.85)^2 + 0.18 * (2 - 0.85)^2 = 0.4875$

$SD[X] = \sqrt{0.4875} = 0.6982$

(2) Define Event H as living in a high-income neighborhood. Define Event W as being white. H and W are independent if there is no segregation. Hence, P(H & W) = P(H)*P(W). Also, P(H') = 1 – P(H).

Joint Probability Table for City X

	H	H'
W	0.0855	0.3645
W'	0.1045	0.4455

(3) In Scenario 2 the expected value of the payout is <u>the same</u> (higher, lower, the same) compared to Scenario 1. In Scenario 2 the standard deviation of the payout is <u>lower</u> (higher, lower, the same) compared to Scenario 1. The payout in Scenario 1 follows a <u>Uniform</u> distribution. The payout in Scenario 2 follows a <u>Normal</u> distribution.

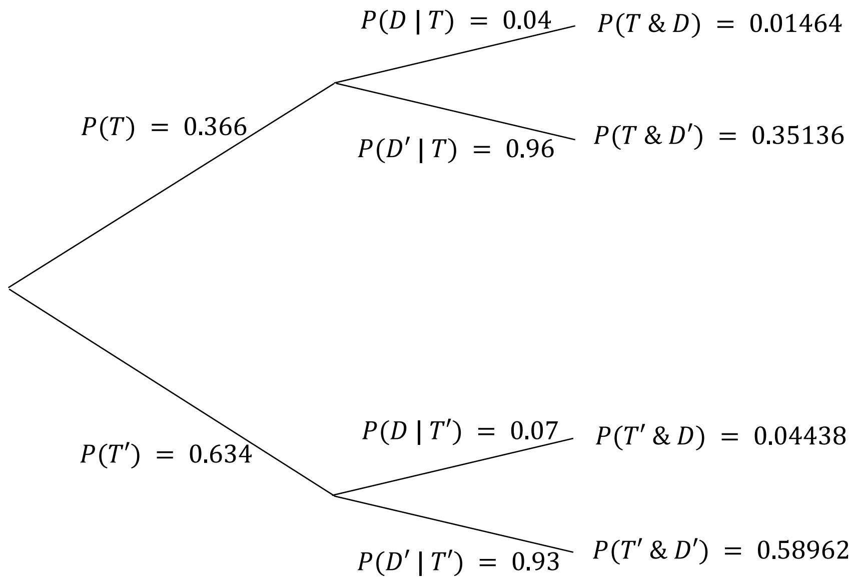
(4) (a) $P(\hat{P} > 0.5 \,|\, n = 1{,}000, p = 0.512) = P\left(Z > \frac{0.5-0.512}{\sqrt{\frac{0.512(1-0.512)}{1{,}000}}}\right) = P(Z > -0.76) = 0.5 + 0.2764 = 0.7764$

(b) The probability would be larger because there will be less sampling error with 10,000 infants born compared to 1,000, and hence a greater chance of being above 50% given that the population parameter is 51.2%.

(c) $P(\hat{P} > 0.52 \,|\, n = 10{,}000, p = 0.512) = P\left(Z > \frac{0.52-0.512}{\sqrt{\frac{0.512(1-0.512)}{10{,}000}}}\right) = P(Z > 1.60) = 0.5 - 0.4452 = 0.0548$

(d) The probability would be larger because there will be more sampling error with 1,000 infants born compared to 10,000, and hence a greater chance of being above 52% even though the population parameter is only 51.2%.

(5)(a) Define the Event T as a borrower attending an institution ranked 1 to 10. Define Event D as default within the first three years.



$P(T \mid D) = ?$

$$P(T \mid D) = \frac{P(T \,\&\, D)}{P(D)}$$

$$P(D) = P(T \,\&\, D) + P(T' \,\&\, D)$$

$$P(D) = 0.01464 + 0.04438 = 0.05902$$

$$P(T \mid D) = \frac{P(T \,\&\, D)}{P(D)} = \frac{0.01464}{0.05902} = 0.24805$$

(b) It would be lower. While it is true that among those attending unranked institutions, the default rate is more than twice as high as at top 10 institutions (10% versus 4%), a very small fraction of borrowers attends unranked institutions (1.9%) compared to top 10 institutions (36.6%). Hence, they would only be a small fraction of the defaults.

(6) (a) $P(U \mid (S \,\&\, 2000)) = \frac{38{,}026}{2{,}361+1{,}116+2{,}165+38{,}026} = 0.8708$

(b) $P(I \mid (C \,\&\, 2010) = \frac{5{,}868}{26{,}298+5{,}868+4{,}320} = 0.1608$ [or could write $P(I \mid (C \,\&\, 2010)$]

(c) $P\big((S \,\&\, C) \mid 2017\big) = \frac{5{,}277}{931{,}394} = 0.0057$

(d) In 2000, there are 114,741 international students from China and India and 5,670 (=4,701+969) went to Canada, which is 4.94%. In 2017, there are 856,458 international students from China and India and 98,777 (=66,161+32,616) went to Canada, which is 11.53%. Hence, Canada has attracted an increasing share of international students from China and India. [*Note:* The question is *not* about what is the Chinese and Indian students' share in Canada. It is about *Canada's* share. The excerpt in the *Supplement* discusses the United States' declining share from 2000 to 2017. This question asks you what is happening to Canada's share. However, the marking TA awarded some partial credit for those that missed the discussion in the excerpt in the *Supplement* and misinterpreted the question.]

(7) (a)

$$P(X > 2 \mid n = 10, p = 0.11) = 1 - P(X = 0) - P(X = 1) - P(X = 2)$$

$$p(0) = \frac{10!}{0!(10-0)!} 0.11^0 (1 - 0.11)^{10-0} = 0.3118$$

$p(1) = \frac{10!}{1!(10-1)!} 0.11^1(1-0.11)^{10-1} = 0.3854$

$p(2) = \frac{10!}{2!(10-2)!} 0.11^2(1-0.11)^{10-2} = 0.2143$

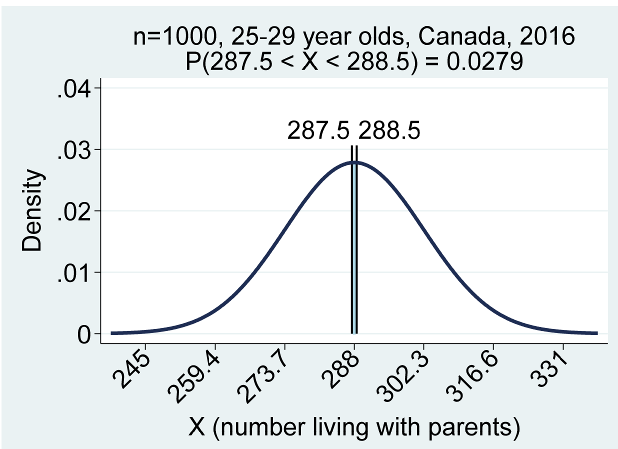
$P(X > 2 \mid n = 10, p = 0.11) = 1 - 0.3118 - 0.3854 - 0.2143 = 0.0885$

(b) The chance that more than 2 have a GPA of 3.0 or higher would be larger in 2002 because of grade inflation: in 1988 only 11% of students had a GPA of 3.0 or more whereas in 2002 19% had such a high GPA. This increases the chance that the sample will include more high GPA students.

(8) $E[X] = np = 1{,}000 * 0.288 = 288$

$SD[X] = \sqrt{np(1-p)} = \sqrt{1{,}000 * 0.288 * (1-0.288)} = 14.32$

$P(287.5 < X < 288.5) = P\left(\frac{287.5-288}{14.32} < Z < \frac{288.5-288}{14.32}\right) = P(-0.0349 < Z < 0.0349) \approx 2 * 0.014 = 0.028$



(9) (a) This is the estimated standard deviation of the sample mean obtained via simulation and it should be close (aside from trivial simulation error) to what theory tells us: $SD[\bar{X}] = \frac{\sigma}{\sqrt{n}} = \frac{17959.62}{\sqrt{36}} = 2993.27$.

(b) If n = 36 were sufficiently large then the simulated sampling distribution of the sample mean (X-bar) would be Normal, according to the Central Limit Theorem (CLT).

$P(Z < 2.33) = 0.99$

$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$

$2.33 = \frac{\bar{X}-124638.3}{17959.62/\sqrt{36}}$

$\bar{X} \approx 131{,}613$

Hence, the 99th percentile of the sampling distribution should be 131,613. Instead, it is 131,991.6 and the discrepancy is because the shape of the sampling distribution is NOT Normal: a sample size of 36 is not sufficiently large in this context given the extreme skew of the ON public sector salary data.