MATH-1230 (DUPRÉ) SPRING 2019 TEST 1 ANSWERS

1. Put the required information in the upper right hand corner of each sheet as specified in the syllabus.

Suppose YOU know that X and Y are unknowns with

$$\mu_X = 70,\ \ \mu_Y = 100,\ \ \sigma_X = 6,\ \ \sigma_Y = 8,\ \ \rho = .4$$

where ρ is the correlation coefficient of X and Y. SLR stands for SIMPLE LINEAR REGRESSION, IRS stands for INDEPENDENT RANDOM SAMPLE, SRS stands for SIMPLE RANDOM SAMPLE.

2. If YOUR FRIEND guesses that the value of X is 78, then what should YOU expect for their squared error?

ANSWER: $(78 - \mu_X)^2 + \sigma_X^2 = 8^2 + 6^2 = 64 + 36 = 100.$

3. What do YOU expect for the value of X^2?

ANSWER: $E(X^2) = \mu_X^2 + \sigma_X^2 = (70)^2 + 6^2 = 4900 + 36 = 4936.$

4. Are X and Y INDEPENDENT? JUSTIFY your answer.

ANSWER: NO, because $\rho \neq 0.$

5. What is the COVARIANCE of X and Y?

ANSWER: $Cov(X, Y) = \rho\sigma_X\sigma_Y = (.4)(6)(8) = 19.2.$

6. What is the VARIANCE of $X - Y$?

ANSWER: $Var(X - Y) = Var(X) + Var(Y) - 2Cov(X, Y) = (6)^2 + (8)^2 - (2)(.4)(6)(8) = 100 - 38.4 = 61.6.$

7. What is $Cov(23, X)$?

ANSWER: ZERO, because the deviation of any constant from its mean is zero.

8. Suppose you observe that X has the value 82. Using SLR, what should you predict for the value of Y?

ANSWER: The standardization of X is $Z_X = (X - \mu_X)/\sigma_X$, so the standard score for a raw score of $x = 82$ is simply $(82-70)/6 = (12)/6 = 2$. This leads to the prediction of the value of Z_Y as being $\rho z_x = (.4)(z_x) = (.4)(2) = (.4)[(82 - 70)/(6)] = .8 = z_y$ for the predicted standard score for Z_Y so this corresponds to the predicted raw score y for Y as

$$y = \mu_Y + \sigma_Y z_y = 100 + (8)(.8) = 100 + (8)(.4)\frac{82 - 70}{6} = 106.4.$$

9. What is the percentage reduction in expected squared error in the predicted value of Y obtained by using SLR as contrasted with simply ignoring any information about X and simply using μ_Y for the predicted value of Y no matter what we find out about X?

ANSWER: $\rho^2 = (.4)^2 = .16 =$**16 percent.**

10. If a fair pair of dice is rolled 3 times, what is the probability that 7 happens exactly once?

ANSWER: Since the results of the different rolls are all independent of each other and as the probability of rolling a 7 is 1/6 on each roll, the count of sevens is governed by the binomial distribution for $n = 3$ **and success rate** $p = 1/6$. **If** X **denotes the number of times we roll a seven, then for** n **independent trials the probability of getting seven to happen** k **times out of** n **is**

$$P(X = k) = C(n, k)p^k(1 - p)^{n-k},$$

so here with $n = 3$ **and** $p = 1/6$

$$P(X = 1) = C(3, 1)(1/6)^1(1 - [1/6])^{3-1} = (3)(1/6)([25]/[36]) = \frac{25}{72}.$$

11. If a fair pair of dice is rolled 24 times in an IRS, then how many times do we expect to get 7?

ANSWER: If X **is the number of successes in** n **trials with either IRS or SRS, then** $E(X) = np$, **where** p **is the success rate. Here we have** $n = 24$ **and** $p = 1/6$, **so**

$$E(X) = (24)\frac{1}{6} = 3.$$

12. A box contains 2 RED blocks and 3 BLUE blocks. If 3 blocks are drawn from the box without replacement in an SRS, what is the probability that NONE are RED?

ANSWER: Drawing without replacement in an SRS the success count is governed by the HYPERGEOMETRIC DISTRIBUTION, so here the population size is $N = 5$, the population of successes size is $R = 2$, for counting RED blocks, and the sample size is $n = 3$. Using X for the success count, we are asking for $P(X = 0)$. In general,

$$P(X = k) = \frac{C(R, k)C(N - R, n - k)}{C(N, n)},$$

so here

$$P(X = 0) = \frac{C(2, 0)C(3, 3)}{C(5, 3)} = \frac{1}{C(5, 3)}.$$

Also,

$$C(5, 3) = C(5, 2) = \frac{(5)(4)}{(2)(1)} = 10,$$

so this means we simply have

$$P(X = 0) = 1/(10).$$

This problem is so simple that it can be worked directly using the elementary methods for dealing with drawing blocks in box problems as discussed in the early problems on the Practice Test Problems document on my website. Getting no red blocks is the same as getting all blue blocks, and as there are only 3 blue blocks of the 5 blocks in the box, this probability is by the ordinary conditional probabilty rule:

$$P(\textbf{ all 3 blue }) = P(1^{st}\textbf{ blue })P(2^{nd}\textbf{ blue }|1^{st}\textbf{ blue })P(3^{rd}\textbf{ blue }|1^{st}\ \&\ 2^{nd}\textbf{ blue })$$
$$= \left(\frac{3}{5}\right)\left(\frac{2}{4}\right)\left(\frac{1}{3}\right) = \frac{1}{10}.$$

13. In the previous problem, what is the VARIANCE in the number of RED blocks drawn?

ANSWER: The distribution for the number of red blocks drawn is hypergeometric, so the expected number of red blocks is

$$E(X) = np = (3)(2/5) = 6/5,$$

and the variance in the number of red blocks drawn is

$$Var(X) = \frac{N - n}{N - 1}np(1 - p) = \frac{N - n}{N - 1}n\left(\frac{R}{N}\right)\left(\frac{N - R}{N}\right)$$
$$= \frac{5 - 3}{5 - 1}(3)\left(\frac{2}{5}\right)\left(\frac{3}{5}\right) = \frac{9}{25}.$$