MATH-1230 (DUPRÉ) SPRING 2019 TEST 2

DIRECTIONS: Put the required information in the upper right hand corner of each sheet as specified in the syllabus. CIRCLE THE CORRECT ANSWERS.

SLR stands for SIMPLE LINEAR REGRESSION, IRS stands for INDEPENDENT RANDOM SAMPLE, SRS stands for SIMPLE RANDOM SAMPLE, PDF stands for Probability Density Function in case of a continuous unknown and Probability Distribution Function in case of a discrete unknown. Z **always denotes the standard normal unknown.**

Suppose that X **and** Y **are INDEPENDENT NORMAL unknowns.** $\bar{W}_n$ **denotes an IRS sample mean for** W **of size** n**, for any unknown** W**. AGAIN,** Z **always denotes the standard normal.**

[1 and 2.] Suppose $\mu_X = 75$ **and** $\sigma_X = 20$**.**

[1.]What is $P(\bar{X}_{16} < 85)$**?**

(A) 0.9772 (B) 0.1587 (C) 0.3413 (D) 3.413 (E) None of the above.
A

ANSWER: Since X **is normal, so is** $\bar{X}_{16}$**. But, even though the mean of** $\bar{X}_{16}$ **is also the same as for** X **itself, its standard deviation is** $\sigma_X/\sqrt{16} = 20/4 = 5$**. Therefore standardizing 85 now results in** $(85 - 75)/5 = 2$**. Therefore**

$$P(\bar{X}_{16} < 85) = P(Z < 2) = 0.9772,$$

so the correct answer choice is A.

[2.] What is a real number x **for which** $P(\bar{X}_4 > x) = 0.975$**?**

(A) 75 +19.6 (B) 75-19.6 (C) 75+39.2 (D) 75-39.2 (E) None of the above.
B

ANSWER: The mean of $\bar{X}_4$ **is again 75, but the standard deviation is** $20/\sqrt{4} = 20/2 = 10$**, so we have**

$$P(\bar{X}_4 \leq x) = 1 - 0.975 = .025,$$

and therefore x **must standardize to** $-z_{0.025} = -1.960$**, which means**

$$\frac{x - 75}{10} = -1.960, \text{ so } x = 75 - 19.6.$$

Thus the correct answer choice is B.

[3 and 4 and 5.] Suppose at a certain busy street intersection one hears a horn honk on average every 4 minutes independent of time of day, in a Poisson process.

[3.] What is the probability one does not hear a horn honk in the first 6 minutes observed given that a honk is not heard in the first 4 minutes observed?

(A) $e^{-3/2}$ **(B)** e^{-1} **(C)** $e^{-6/4}$ **(D)** $e^{-1/2}$ **(E) None of the above.**
D

ANSWER: A Poisson process has no memory, so given no honks in the first four minutes means as far as the the first six minutes is concerned it is the same as if we just started observing after 4 minutes and listen for 2 more minutes. But you only expect .5 honks in 2 minutes. Thus, the probability is the same as the probability of expecting 1/2 but hearing 0 for the Poisson distribution and this is simply $e^{-1/2}$**. The correct answer choice is simply D.**

[4.] What is the probabilty that you have to wait more than 20 minutes to hear the fifth honk?

(A) 0.6288 (B) 0.6160 (C) 0.4405 (D) 0.7851 (E) None of the above.
C

ANSWER: Having to wait more than 20 minutes to hear the fifth honk is equivalent to hearing AT MOST 4 honks within the first 20 minutes. However, in 20 minutes you expect to hear 5 honks since they happen on average every 4 minutes. Therefore the probability we seek is simply the value of the Poisson cdf evaluated at $k = 4$ **for the case where** $\mu = 5$**. From the table of Poisson probabilities we easily see this is 0.4405, so the correct answer choice is C.**

[5.] What is the probability that during a 20 minute timespan observed you hear exactly 4 honks?

(A) 0.1563 (B) 0.1953 (C) 0.1708 (D) 0.1755 (E) None of the above.
D

ANSWER: If T **denotes the total number of honks heard in 20 minutes, as we expect 5, we have** $E(T) = \mu = 5,$ **and from the Poisson cdf table for** $\mu = 5,$

$$P(T = 4) = P(T \leq 4) - P(T \leq 3) = 0.4405 - 0.2650 = 0.1755.$$

The correct answer choice is therefore D.

[6 and 7.] Suppose that W is a continuous unknown with PDF satisfying

$$f_W(w) = 4w^3, \ 0 < w < 1.$$

[6.]What is the expected value of W?

(A) 1/5 (B) 4/5 (C) 1 (D) 16/5 (E) None of the above.
B

ANSWER: We need to use our integral formula for expected value. Remember, for any continuous function $g : \mathbb{R} \longrightarrow \mathbb{R}$ from the reals to the reals we have

$$E(g(W)) = \int_{\mathbb{R}} g(w) f_W(w) dw$$

Moreover, as

$$\int_0^1 f_W(w)dw = 4 \cdot \frac{1}{4} w^4|_0^1 = 1,$$

it follows that $f_W(w) = 0$ unless $0 \leq w \leq 1$, so all integrals over the whole real line having f_W as a factor in the integrand can be replaced by integrals from 0 to 1. In particular, for any positive integer n, we have

$$E(W^n) = \int_0^1 w^n \cdot 4w^3 dw = 4 \int_0^1 w^{n+3} dw = \frac{4}{n+4}.$$

In particular, $E(W) = 4/5$ using $n = 1$. Therefore the correct answer choice is B.

[7.] What is the expected value of W^2?

(A) 2/3 (B) 16/25 (C) 34/75 (D) 1/4 (E) None of the above.
A

ANSWER: Referring to the previous answer, we have using $n = 2$, that

$$E(W^2) = 4/6 = 2/3,$$

so the correct answer choice is A.

[8 and 9.] Suppose that W is continuous and only known to have maximum value 12 and minimum value 4.

[8.] What is $P(5 < W < 9|W > 6)$?

(A) 1/3 (B) 1/4 (C) 2/3 (D) 1/2 (E) None of the above.
D

ANSWER: The statement "$5 < W < 9$ and $W > 6$" is the same as $6 < W < 9$. Using the formula for conditional probability, we know for any statements A and B we have

$$P(A|B) = \frac{P(A \textbf{ and } B)}{P(B)} \textbf{ so } P(5 < W < 9|W > 6) = \frac{P(6 < W < 9)}{P(W > 6)}.$$

Sine we only know the minimum and maximum values of W, it follows that W must be uniformly distributed between 4 and 12, so the numerator probability is 3/8 while the denominator probability is 6/8. The quotient is therefore 3/6=1/2, so the correct answer choice is D.

[9.] What is the expected value of W?

(A) 8 (B) 6 (C) 2 (D) 0 (E) None of the above.
A

ANSWER: The expected value of W is the average of its maximum and minimum values since it is uniformly distributed. The average of 4 and 12 is obviously 8, so the correct answer choice is A.

[10 and 11.] Suppose that $\bar{x}$ is an observed value of $\bar{X}_{100}$, keeping in mind that X is normal.

[10.] If $\sigma_X = 10$, what is the MARGIN OF ERROR in the 99 percent Confidence Interval for μ_X?

(A) 2.326 (B) 23.26 (C) 2.576 (D) 25.76 (E) None of the above.
C

ANSWER: Remember, the formula for the margin of error M is simply

$$M = (\textbf{table stat})(SE),$$

where SE denotes the STANDARD ERROR, that is the standard deviation of the estimator. In this case our estimator is $\bar{X}_{100}$ whose standard deviation is

$$SE = \frac{\sigma_X}{\sqrt{n}} = \frac{10}{\sqrt{100}} = 1,$$

and as $\bar{X}_{100}$ is normal because X is, the appropriate statistic is the critical value for the standard normal distribution cutting a right tail of area .005. From the $t-$table of critical values with infinity degrees of freedom this critical value is 2.576, so the correct answer choice is C.

[11.] If σ_X is unknown and we use an observed sample variance $s^2 = 81$, for a sample of size $n = 9$, then what is the MARGIN OF ERROR in the 99 percent Confidence Interval for μ_X based on this small sample?

(A) (3)(2.576) (B) (3)(3.355) (C) (9)(3.250) (D) (1/9)(3.169) (E) None of the above.
B

ANSWER: Now we are having to replace σ_X whose value we do not know with the sample standard deviation $s = 9$. This means the appropriate table statistic we need is the critical value for the $t-$ distribution with 8 degrees of freedom cutting off a right tail of area .005. From the table, we see this critical value is 3.355. On the other hand, our standard error is now, though only an estimate, given by

$$SE = \frac{9}{\sqrt{9}} = 3.$$

Our margin of error is therefore (3)(3.355), so the correct answer choice is B.

[12 and 13.] Suppose a box filled with beads each of which is either red or white. They are so small that without looking closely, the box appears to be filled with pink beads. Assume the box contains so many beads, that removing 300 beads has a negligible effect on the overall proportion of red beads and we do not know the total number of beads in the box.

[12.] If the true proportion of red beads in the box is 1/4, and if T denotes the number of red beads in an SRS of size 300, what is the best approximate value of the probability that the number of red beads in the sample is at most 60?

(A) $P(Z \geq 3/4)$ **(B)** $P(Z \leq -31/15)$ **(C)** $P(Z \leq -2)$ **(D)** $P(Z \leq -29/15)$ **(E) None of the above.**
D

ANSWER: Since the box has so many beads that taking a few hundred out does not change the overall proportion of red beads in any significant way, we can assume that SRS is the same as IRS for samples of size $n = 300$. This means that even though technically, T has the hypergeometric distribution, we can assume the T has the binomial distribution for calculation purposes, with $n = 300$ as the number of independent trials and $p = 0.25$ as the success rate. Then the expected number of red beads in our sample is $\mu = np = 75$. As n is too large for our binomial tables, we must use the normal approximation which is certainly valid here as we expect 75 successes and 225 failures, both way beyond the required 5. The Variance of T is

$$Var(T) = npq, \text{ where as usual, } q = 1 - p \text{ so here } n = 300, \ p = 1/4, \ q = 3/4,$$

and thus

$$Var(T) = (300)(1/4)(3/4) = 3^2 \cdot 10^2/4^2, \text{ so } \sigma_T = 3 \cdot 10/4 = 7.5.$$

Therefore

$$\mu_T = 75 \text{ and } \sigma_T = 7.5.$$

To get the most accurate normal approximation, we must employ the continuity correction. Thus, thinking of our normal random variable X as having mean 75 and standard deviation 7.5, we have

$P(T \leq 60)$ **is best approximated as** $P(X \leq 60.5)$. **But,**

$$P(X \leq 60.5) = P(Z \leq [60.5 - 75]/(7.5) = P(Z \leq -(14.5)(7.5)) = P(Z \leq -29/15).$$

The correct answer choice is therefore D.

[13.] If the true proportion of red beads in the box is 1/4, what is the best approximation to the probability the number of red beads in an SRS of size 300 is at most 75

(A) 1/2 (B) 0.5279 (C) 0.4721 (D) 3/4 (E) None of the above.
B

ANSWER: Using the continuity correction, $P(Z \leq (0.5)/(7.5)) = P(Z \leq 0.07) = 0.5279$, **as (0.5)/(7.5)=1/15 which is 0.07333... Choose B.**

[14 through 20.] In the following problems choose the correct name of the distribution needed for calculating the probability indicated.

[14 through 20. ANSWERS: A B C E D F G

[14.] That 3 observations of X in an IRS of size 5 are above average.

(A) BINOMIAL (B) NORMAL (C) CHI SQUARE (D) HYPERGEOMETRIC (E) EXPONENTIAL (F) F-DISTRIBUTION (G) None of the above.

ANSWER: This is like asking the probability of 3 heads in five flips of a fair coin, so the distribution is binomial.

[15.] That $|X - Y| \leq \sqrt{\sigma_X^2 + \sigma_Y^2}$.

(A) BINOMIAL (B) NORMAL (C) CHI SQUARE (D) HYPERGEOMETRIC (E) EXPONENTIAL (F) F-DISTRIBUTION (G) None of the above.

ANSWER: As X and Y are independent normals, so is $X - Y$, therefore this is a simple application of the standard normal distribution.

[16.] That the sample variance of an IRS of size 8 for X is more than $2\sigma_X^2$.

(A) BINOMIAL (B) NORMAL (C) CHI SQUARE (D) HYPERGEOMETRIC (E) EXPONENTIAL (F) F-DISTRIBUTION (G) None of the above.

ANSWER: X is normal so $8 \cdot S^2/\sigma_X^2$ is chi-square for 8 degrees of freedom.

[17.] Of having to wait more than 10 minutes for a trollley when told that on average they arrive every 5 minutes.

(A) BINOMIAL (B) NORMAL (C) CHI SQUARE (D) HYPERGEOMETRIC (E) EXPONENTIAL (F) F-DISTRIBUTION (G) None of the above.

ANSWER: This is the waiting time in a Poisson process, so the distribution is the exponential distribution.

[18.] Of drawing 3 aces in a five card poker hand.

(A) BINOMIAL (B) NORMAL (C) CHI SQUARE (D) HYPERGEOMETRIC (E) EXPONENTIAL (F) F-DISTRIBUTION (G) None of the above.

ANSWER: A poker hand is an SRS from a population of size $N = 52$, so we use the hypergeometric distribution.

[19.] That S_X^2/S_Y^2 is less than twice σ_X^2/σ_Y^2 with both IRS samples of size 5.

(A) BINOMIAL (B) NORMAL (C) CHI SQUARE (D) HYPERGEOMETRIC (E) EXPONENTIAL (F) F-DISTRIBUTION (G) None of the above.

ANSWER: Since both numerator and denominator have 5 degrees of freedom, this is obviously $(\sigma_X^2/\sigma_Y^2) \cdot F(5, 5)$. Choose F.

[20.] Of engine failure due to excessive wear within half the time expected.

(A) BINOMIAL (B) NORMAL (C) CHI SQUARE (D) HYPERGEOMETRIC (E) EXPONENTIAL (F) F-DISTRIBUTION (G) None of the above.

ANSWER: You might be tempted to guess this is a Poisson process, but waiting for something to wear out is not memoryless. The longer an engine runs, the more likely it is to fail, so this is governed by a distribution which is not in our list. It is more complicated than the Poisson process, but does have a useful analysis. We are not covering it. However, you always need to be aware of the assumptions required for any distribution you might be inclined to use so that you do not arrive at erroneous conclusions. Choose G.

[14 through 20. ANSWERS: A B C E D F G