MATH-1110 (DUPRÉ) SPRING 2011 TAKE HOME TEST 1 ANSWERS

FIRST: PRINT YOUR LAST NAME IN LARGE CAPITAL LETTERS ON THE UPPER RIGHT CORNER OF EACH SHEET TURNED IN.

SECOND: PRINT YOUR FIRST NAME IN CAPITAL LETTERS DIRECTLY UNDERNEATH YOUR LAST NAME ON EACH SHEET TURNED IN.

THIRD: WRITE YOUR SPRING 2011 MATH-1110 SECTION NUMBER DIRECTLY UNDERNEATH YOU FIRST NAME ON EACH SHEET TURNED IN.

FOURTH: THERE ARE TWENTY FIVE QUESTIONS AND EACH IS WORTH 4 POINTS. WRITE ALL YOUR ANSWERS NEATLY IN THE SPACE PROVIDED UNDER EACH QUESTION. NEATNESS COUNTS. IF I CANNOT READ IT WITHOUT STRAINING MY EYES YOU GET NO CREDIT.

Suppose that the fish in my pond have mean length 7 inches with a standard deviation of 1.3 inches, and mean weight 2.2 pounds with a standard deviation of 0.4 pounds. Suppose that the correlation between length and weight is $\rho = .7$ Suppose that a fish (henceforth to be referred to as "the fish") is taken from my pond with length X and weight Y.

1. What is the optimal guess for the length of the fish, that is, what is $E(X)$?

$E(X) = \mu_X = 7.$

FINAL ANSWER: 7

2. What is $E(Y)$?

$E(Y) = \mu_Y = 2.2.$

FINAL ANSWER: 2.2

3. What is the expected squared error if you guess the length of the fish to be $E(X)$?

The expected squared error is $E([X-7)^2] = E([X-\mu_X]^2) = Var(X) = \sigma_X^2 = (1.3)^2 = 1.69.$

FINAL ANSWER: 1.69

4. What is the expected squared error if you guess the weight of the fish to be $E(Y)$?

The expected squared error is $Var(Y) = \sigma_Y^2 = (0.4)^2 = 0.16.$

FINAL ANSWER: 0.16

5. If you have the information that the fish is actually 9.6 inches long, then what is the optimal guess for the weight of the fish using that information?

Since the standard deviation in length is 1.3 inches, a fish which is 9.6 inches long is 2 standard deviations above the mean, since the mean is only 7, and therefore, if the correlation between length and weight were perfect in which case, we would have ρ, the correlation coefficient, equal to one, then we would guess that the weight is also 2 standard deviations above the mean, which would be 0.8 pounds above the mean of 2.2, or 3 pounds. However, the correlation is not perfect, it is $\rho = .7$, so we only guess the weight to be $2 \cdot (0.7) = 1.4$ standard deviations above the mean. As the standard deviation for weight is 0.4 pounds, 1.4 standard deviations is $(0.4)(1.4) = 0.56$, so we should guess the fish to be only 0.56 pounds above the average of 2.2 poounds, which is a weight of 2.76 pounds.

FINAL ANSWER: 2.76

6. What is $E(Y|X = 9.6)$?

The expression $E(Y|X = 9.6)$ is the expression for the expected value of Y given the information that $X = 9.6$. Since the expected value using the information available is the optimal guess, we know that the answer here is exactly the same as for the previous problem.

FINAL ANSWER: 2.76

7. What is the expected squared error in your guess of the weight of a 9.6 inch fish when you guess $E(Y|X = 9.6)$?

When you guess $E(Y|X = 9.6)$, for the weight of a 9.6 inch fish, you are using linear regression in which case your expected squared error is now

$$E([Y - 2.76]^2)|X = 9.6) = (1 - \rho^2) \cdot \sigma_Y^2 = (1 - [0.7]^2) \cdot (0.4)^2 = (.51)(.16) = 0.0816.$$

FINAL ANSWER: 0.0816

Suppose that a wildlife biologist is studying the relationship between length and weight of Pacific salmon. In a sample of 100 fish, the sample mean length of the fish was 27 inches with a sample standard deviation $s = 7.2$ inches and the sample mean weight was 8.3 pounds with a sample standard deviation of $s = 3.8$ pounds, and additionally, the sample correlation coefficient between length and weight is $r = .8$. The wildlife biologist is told that a Pacific salmon of length X and weight Y was caught by a local fisherman (henceforth referred to as "the fish"). Let D denote the statement of the wildlife biologist's sample data.

8. What should the wildlife biologist guess for the length of the fish given his sample data, that is, what is $E(X|D)$?

Before we look at the sample data, the sample mean is itself an unknown, denoted $\bar{X}$ and $E(\bar{X}) = E(X) = \mu_X$, consequently, since the sample mean length of the fish in the sample is 27 inches, that is the optimal guess, given the sample data D. That is to say simply, $E(X|D) = 27$. In general, based only on a sample mean, the optimal guess for an unknown is that sample mean.

FINAL ANSWER: 27

9. What is $E(Y|D)$?

Likewise as in the previous problem, $E(Y|D) = 8.3$, since the mean weight of the fish in the sample is 8.3.

FINAL ANSWER: 8.3

10. If the wildlife biologist finds that the length of the fish is 34.2 inches long, using this information along with the sample data, what should the wildlife biologist guess for the weight of the fish, that is, what is $E(Y|(X = 34.2)\&D)$?

A fish which is 34.2 inches long is 7.2 inches above the sample mean. If we knew how many standard deviations above the mean that is and if we knew the correlation between length and weight, we would use linear regression like in the previous problem to guess the weight. Our problem know is to begin by guessing the standard deviations and the correlation coefficient. Now, the sample variance is expected to be the true variance, so we use the sample standard deviations as our guesses for the true standard deviations, and the sample correlation coefficient r is expected to be the true correlation coefficient ρ. That it, we use r in place of the unknown value of ρ, in our calculations. Since our fish is one sample standard deviation above the sample mean and the sample correlation coeffiicient is $r = .8$, we would guess that our fish is 0.8 weight sample standard deviations above the sample mean in weight. As the sample standard deviation for weight is 3.8, the fish is guessed to be $(0.8)(3.8) = 3.04$ pounds above the sample mean weight of 8.3 pounds, which would be finally 11.34 pounds.

FINAL ANSWER: 11.34

Suppose that a standard deck of cards is used to play a game of poker. There are four players A, B, C, D and a professional dealer who does not play. The dealer deals each player 5 cards face down.

11. What is the probability that player A receives three aces?

Since there are 52 nCr 5 $= 2598960$ five card hands when five cards are dealt from a standard deck of 52 cards, we need to find the number of ways to get three aces. There is only one way to get four aces in four cards, but we have a five card hand, and there are then 48 ways to choose a fifth card, so there are 48 ways to get a five card hand with four aces. To get a five card hand with exactly three aces, there are 4 nCr 3 $= 4$ ways to choose three of the four aces in the deck, and then 48 nCr 2 $= 1128$ ways to choose the remaining two cards for your five card hand, so there are $(4)(1128) = 4512$ ways to make a hand with exactly three aces. This means the number of hands with three aces is $48 + 4512 = 4560$. Therefore, the probability of getting three aces (or more) is $4560/2598960 = 0.001754548$ or about 0.00175, to three significant digits. Notice that if you interpret getting three aces to mean exactly three aces, your answer would be $4512/2598960 = 0.001736079$ or about 0.00174, which is only off by less than one percent.

WARNING: There is a subtle pitfall here in counting. You might be tempted to say step one choose three aces and step two choose two other cards from the remaining 49 cards in the deck in order to count the number of ways to get at least three aces. Keep in mind that when using the Multiplication Principle, that your steps should have the property that any decision change at any step will change the final result of going through the steps. The two steps here do not have that property. If you first get all the aces except the ace of hearts and then on the second step get the ace of hearts and the king of hearts, the result is the same as if on the first you get all the aces except the ace of diamonds and on the second step get the ace of diamonds and te king of hearts. Either way, you end up with all four aces and the king of hearts. This would violate the multiplication principle. The first step can be done in 4 nCr 3 $= 4$ ways, and the second can be done in 49 nCr 2 $= 1176$ ways, and that product is 4704 which is more than the number of ways to get at least three aces which we know already is only 4560. The discrepency is caused by the double counting in this last stepwise procedure.

FINAL ANSWER: 0.00175

12. What is the probability that player B receives 3 kings and 2 aces?

Now we need to know how many ways to make a hand with 3 kings and 2 aces. We can make such a hand in two steps. First get the kings and second get the aces. There are $4\ nCr\ 3 = 4$ ways to get three of the four kings, and $4\ nCr\ 2 = 6$ ways to get the two aces, so there are $(4)(6) = 24$ ways to get 3 kings and two aces, and therefore as there are 2598960 possible five card hands, the probability of being dealt 3 kings and two aces is $24/2598960 = 0.00000923446301$ or about 0.00000923, to three significant digits.

FINAL ANSWER: 0.00000923

13. What is the probability that player B receives 3 kings and 2 aces given that player A receives exactly one ace?

If one ace is removed from consideration because another player is known to have it, then the number of possible five cards hands that player B can get is only $51\ nCr5 = 2349060$. But, from the 51 card deck with only three aces, there is only $3\ nCr\ 2 = 3$ ways to get 2 aces and $4\ nCr\ 3$ ways to get three kings, and therefore only $(4)(3) = 12$ ways for player B to get three kings and two aces given that player A receives exactly one ace. Therefore the probability that player B gets three kings and 2 aces is $12/2349060 = 0.000005108426349$ or about 0.00000511, to three significant digits, or slightly over half as likely as when we do not know what other players have.

14. What is the probability that player B receives 3 kings and two aces given that player A receives exactly one ace and player C receives exactly one king?

Now, we are effectively removing two cards from the deck as far as what is possible for player B. The number of possible hands is $50\ nCr\ 5 = 2118760$, and there is only one way player B can get three kings and only 3 ways to get two aces, so there are only 3 ways for player B to get 3 kings and 2 aces now, so his probability of getting 3 kings and 2 aces is only $3/2118760 = 0.000001415922521$, or about 0.00000142, to three significant digits.

FINAL ANSWER: 0.00000142

15. What is the probability that player A receives all diamonds?

To make a hand with all diamonds, you need to choose all 5 cards from the 13 diamonds in the deck, so there are $13\ nCr\ 5 = 1287$ ways to make a hand with all diamonds. since there are 2598960 possible five card hands, this makes the probability of getting all diamonds is $1287/2598960 = 0.0004951980792$, or about .000495, to three significant digits.

FINAL ANSWER: 0.000495

16. What is the probability that player B receives all cards of the same suit?

The probability that all cards are the same suit is the sum of the probabilities for each of the four suits which of course are the same as for the diamonds, so the probability is the four times the answer to the previous problem, or $(4)(1287)/2598960 = 0.0019807923$, or about 0.00198, to three significant digits.

FINAL ANSWER: 0.00198

17. What is the probability that player A receives 5 diamonds given that player C receives 5 diamonds?

Now, there are 5 diamonds removed from being chosen for player A. It is as if he is being dealt a hand of 5 cards from a deck of 47 cards with only 8 diamonds. Therefore his chance of getting 5 diamonds is $(8\ nCr\ 5)/(47\ nCr\ 5) = 56/1533939 = 0.00003650731874$, or about 0.0000365, to three significant digits.

FINAL ANSWER: 0.0000365

Suppose that we are writing a string of letters in line (that is a "word") taken from the alphabet A, B, C, D, K, L, M.

18. How many words of length 5 letters are possible?

Each letter position in the word can have any of the 7 letters in our alphabet, so the number of possible 5 letter words is $7^5 = 16807$.

FINAL ANSWER: 16807

19. How many 5 letter words are possible if all the letters in the word must be different?

If all the letters must be different, then a five letter word amounts to an arrangement of five of the letters taken from the set of 7 letters available which is $7\ nPr\ 5 = 2520$.

FINAL ANSWER: 2520

20. How many 10 letter words are possible which have three A's, four B's and three K's?

Since you cannot tell how the A's are arranged among themselves-it does not matter, and likewise for the B's and K's, the answer is

$$\frac{10!}{(3!)(4!)(3!)} = 4200.$$

FINAL ANSWER: 4200

Suppose that X **and** Y **are unknowns and that** $\mu_X = 8$, **that** $\sigma_X = 4$, **that** $\mu_Y = 9$, **that** $\sigma_Y = 5$, **and that the correlation of** X **with** Y **is** $\rho = .7$.

21. What is $Cov(X, Y)$?

Since $Cov(X, Y) = \rho\sigma_X\sigma_Y$, we can use what is given to find $Cov(X, Y) = (.7)(4)(5) = 14$.

FINAL ANSWER: 14

22. What is $E(X)E(Y)$?

This is simply the product of the means since $E(X) = \mu_X$, and $E(Y) = \mu_Y$, therefore $E(X)E(Y) = (8)(9) = 72$.

FINAL ANSWER: 72

23. What is $E(XY)$?

We know in general we know that

$$Cov(X, Y) = E(XY) - E(X)E(Y) = E(XY) - \mu_X\mu_Y,$$

so

$$E(XY) = \mu_X\mu_Y + Cov(X, Y),$$

which is here $E(XY) = 72 + 14 = 86$.

FINAL ANSWER: 86

24. What is $E[(X-5)(Y-4)]$?

We just use algebra to simplify what is "inside" the expectation,

$$E[\text{ simplify what is in here }],$$

and then use the rules of expectation.

$$(X-5)(Y-4) = XY - 5Y - 4X + 20,$$

therefore

$$E[(X-5)(Y-4)] = E[XY - 5Y - 4X + 20] = 86 - (5)(9) - (4)(8) + 20$$

$$= 86 - 45 - 32 + 20 = 29.$$

Alternately, we can use

$$Cov(X-5, Y-4) = Cov(X,Y) = 14,$$

since adding or subtracting constants does nothing to variance or covariance. Then use the same method as the previous problem, since

$$E(X-5) = 8 - 5 = 3,$$

and

$$E(Y-4) = 9 - 4 = 5,$$

therefore

$$E[(X-5)(Y-4)] = E(X-5)E(Y-4) + Cov(X-5, Y-4)$$

$$= (3)(5) + 14 = 15 + 14 = 29.$$

FINAL ANSWER: 29

25. What is the standard deviation of $X - Y$?

First, remember standard deviation is the square root of variance and that $Var(X) = Cov(X, X)$ for any unknown X, so $Var(X - Y) = Cov(X - Y, X - Y)$.

We have to know that Cov obeys rules like expectation, in each of its "variables", that is, for any unknowns W, X, Y and any constant c, it is the case that

$Cov(X, Y) = Cov(Y, X)$,
$Cov(W + X, Y) = Cov(W, Y) + Cov(X, Y)$,
$Cov(c \cdot X, Y) = c \cdot Cov(X, Y)$.

You should notice that these rules are just like the usual rules for ordinary algebra with numbers

$ab = ba$,
$(a + b)c = ac + bc$,
$(ca)b = c(ab)$.

The same algebra that shows

$$(a \pm b)^2 = (a + b)(a + b) = a^2 + b^2 \pm 2ab,$$

then shows that

$$Cov(X + Y, X + Y) = Cov(X, X) + C0v(Y, Y) \pm 2Cov(X, Y),$$

or

$$Var(X \pm Y) = Var(X) + Var(Y) \pm 2Cov(X, Y).$$

For our situation here, we know that $Var(X) = 4^2 = 16$, and $Var(Y) = 5^2 = 25$, so

$$Var(X - Y) = 16 + 25 - (2)(14) = 16 + 25 - 28 = 16 - 3 = 13.$$

Therefore, the standard deviation of $X - Y$ is $\sigma = \sqrt{13} = 3.605551275$, or about 3.61, to three significant digits.

FINAL ANSWER: 3.61