MATH-1110 (DUPRÉ) PRACTICE TEST PROBLEM ANSWERS

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

SECOND: PRINT YOUR FIRST NAME IN CAPITAL LETTERS DIRECTLY UNDERNEATH YOUR LAST NAME.

THIRD: WRITE YOUR FALL 2010 MATH-1110 LAB DAY DIRECTLY UNDERNEATH YOU FIRST NAME.

STANDARD INFORMATION: A **Standard Dice** has six faces and each face has a number of spots, that number being a positive whole number no more than six. The total number of spots on any pair of opposite faces is seven. Since there are exactly three pair of opposite faces, the total number of spots on a dice is twenty one. Therefore if we think of the spots as distributed evenly over the faces, there are on average 3.5 spots per face. A **Standard Deck of Cards** has 52 cards, four suits of 13 cards (denominations) each. The four suits are spades ($\spadesuit$), hearts($\heartsuit$), diamonds($\diamondsuit$), and clubs($\clubsuit$). Each suit has 3 face cards: Jack, Queen, King, and an Ace. In each suit, the cards which are not face cards each have a number of spots in the shape of the particular suit, the number of spots being a whole number no more than ten, the card with a single spot being the Ace of that suit. For example, the card with four spots each in the shape of a diamond is called "the four of diamonds($4\diamondsuit$)", whereas the card with the face labelled with "J" and a spade shaped spot is "Jack of spades($J\spadesuit$)". In many card games, an Ace can count as a denomination value of one or as a denomination higher King as a player desires. A Jack has denomination value eleven, a Queen has denomination value twelve, and a King has denomination value 13, in many card games. In the game of Black Jack or Twentyone, all face cards are given denomination value ten and all aces have denomination value eleven or one as desired by the player. A **Standard (American) Roulette Wheel** has a spinner with 38 slots which spins in a large bowl. A ball is sent rolling in opposite direction to the spin of the wheel near the upper rim of the bowl and as both the spinning wheel and ball slow down, the ball falls into one of the 38 slots. Two of the slots are colored green, one labelled zero (0) and the other labelled with two zeros (00), and referred to as "double zero". The remaining slots are each colored either red or black and numbered with the positive whole numbers no more than 36. At the roulette table a player can bet on zero or on double zero or on a specific positive whole number, or on even or on odd or on red or on black or on one through eighteen, or on one through twelve, or on 13 through 24, or on 25 through 36. There are therefore many possibilities for placing bets at the roulette wheel. The game of craps is played on a large oblong table with rounded ends and with vertical walls around the edges. The player who rolls the dice is called "the shooter" and must toss a pair of standard dice so as to hit the table's horizontal surface and bounce off the wall at the opposite end of the table back on the horizontal surface where it finally comes to rest. If any side is touching a vertical wall or is tilted off horizontal or goes off the table, then it does not count and the pair of dice must be tossed again-it is a "do over". Notice that it is virtually impossible to tell what faces will come up when a pair of dice is thrown, or which card will come up when a card is taken off the top of a shuffled deck, or what slot the ball will land on when the roulette wheel is spun. All the processes obey the laws of physics which are completely deterministic. However, the circumstances are guaranteed to make the use of physics to predict the outcomes virtually useless, as there is no practical way to keep track of the information required. Thus, the so-called "random" processes are really "randumb" processes, as they depend on reducing our information below what can be used to make any effective prediction. On the other hand, it is known that a very few people with practice have learned to throw the dice so as to have a fair amount of control over the outcomes and as of this time the major casinos do not admit that as a possibility so these few players have a tremendous advantage at the crap table.

1. Suppose that Sam guesses that his Master Charge card balance is 1450 dollars and his Visa card balance is 2500 dollars. What should he guess for the total he owes on both credit cards in order to be consistent?

ANSWER
$1450 + 2500 = 3950$ dollars.

2. If a dice is in a box where you cannot see it and I look into the box and see it and tell you it is sitting with one of its faces on the floor of the box and one face on top-it is not tilted, then based only on this information (whether or not you believe it) what should you guess is the probability that the face on top has 4 spots?

ANSWER
$\frac{1}{6}$

3. If a dice is in a box where you cannot see it and I look into the box and see it and tell you it is sitting with one of its faces on the floor of the box and one face on top-it is not tilted, then based only on this information (whether or not you believe it) what should you guess is the probability that the face on top has an even number of spots?

ANSWER
$\frac{3}{6} = \frac{1}{2}$

4. If a dice is in a box where you cannot see it and I look into the box and see it and tell you it is sitting with one of its faces on the floor of the box and one face on top-it is not tilted, then based only on this information (whether or not you believe it) what should you guess is the probability that the face on top has an odd number of spots?

ANSWER
$\frac{3}{6} = \frac{1}{2}$

5. If a dice is in a box where you cannot see it and I look into the box and see it and tell you it is sitting with one of its faces on the floor of the box and one face on top-it is not tilted and the top face has an even number of spots, then based only on this information (whether or not you believe it) what should you guess is the probability that the face on top has 4 spots?

ANSWER
$\frac{1}{3}$

6. If a dice is in a box where you cannot see it and I look into the box and see it and tell you it is sitting with one of its faces on the floor of the box and one face on top-it is not tilted and the top face has an even number of spots, then based only on this information (whether or not you believe it) what should you guess is the probability that the face on top has 5 spots?

ANSWER
0

7. If a dice is in a box where you cannot see it and I look into the box and see it and tell you it is sitting with one of its faces on the floor of the box and one face on top-it is not tilted and the top face has an even number of spots, then based only on this information (whether or not you believe it) what should you guess is the number of spots on the top face?

ANSWER

The only possible values here are 2,4,6 and all three of these values are equally likely, so we should guess the average of these three numbers which is 4.

4

8. If a dice is in a box where you cannot see it and I look into the box and see it and tell you it is sitting with one of its faces on the floor of the box and one face on top-it is not tilted and the top face has an odd number of spots, then based only on this information (whether or not you believe it) what should you guess is the number of spots on the top face?

ANSWER

The only possible values here are 1,3,5 and all three of these values are equally likely, so we should guess the average of these three numbers which is 3.

3

9. If X is a positive whole number that I have chosen and you think X is three times as likely to be even as odd, then what should you think is the probability that X is odd?

ANSWER

It is the same as if you have a deck of cards and each has either "even" or "odd" written on it, what is the chance of drawing a card that has "odd" written on it, given that there are three times as many cards with "even" written on them as with "odd" written on them. The simplest example is a deck of 4 cards where only one has "odd" written on it and three each have "even" written on it. From such a deck, the chance the top card has "odd" written on it is clearly 1/4.

$\frac{1}{4}$

10. If X is a positive whole number with $X \leq 6$ that I have chosen and you think X is three times as likely to be even as odd, and if K is a symbol that stands for this information, then what is $E(X|K)$?

ANSWER

Let A denote the statement that X is odd and let B be the statement that X is even. Then just as above, we know

$$P(A|K) = \frac{1}{4}$$

and therefore

$$P(B|K) = 1 - P(A|K) = \frac{3}{4}.$$

We also know that no odd number is any more likely than any other odd number, so

$$E(X|A\&K) = 3,$$

whereas since no even number is any more likely than any other even number,

$$E(X|B\&K) = 4.$$

Now we just apply our general formula that applies whenever there are several statements of which exactly on must be true. In this case we are dealing with, we know either A or B must be true and only one of these two statements can be true. Therefore

$$E(X|K) = E(X|A\&K)P(A|K) + E(X|B\&K)P(B|K) = (3)(\frac{1}{4}) + (4)(\frac{3}{4}) = \frac{15}{4}.$$

$\frac{15}{4}$

11. If fourteen cards are dealt from the top of a well shuffled deck of cards one after another without replacement, what is the chance that the top card is a heart?

ANSWER

Since all four suits have the same number of cards, that means all four suits are equally likely to be the suit of the card on top. This means the probability the top card is a heart is 1/4.

$\frac{1}{4}$

12. If fourteen cards are dealt from the top of a well shuffled deck of cards one after another without replacement, what is the chance that the top card is a diamond?

ANSWER

Since all four suits have the same number of cards, that means all four suits are equally likely to be the suit of the card on top. This means the probability the top card is a diamond is 1/4.

$\frac{1}{4}$

13. If fourteen cards are dealt from the top of a well shuffled deck of cards one after another without replacement, what is the chance that the second card is a heart?

ANSWER

Since all four suits have the same number of cards, that means all four suits are equally likely to be the suit of the card underneath the top card. This means the probability the second card is a heart is 1/4.

$\frac{1}{4}$

14. If fourteen cards are dealt from the top of a well shuffled deck of cards one after another without replacement, what is the chance that the third card is a heart?

ANSWER

Since all four suits have the same number of cards, that means all four suits are equally likely to be the suit of the third card from the top. This means the probability the third card is a heart is 1/4.

$\frac{1}{4}$

15. If fourteen cards are dealt from the top of a well shuffled deck of cards one after another without replacement, what is the chance that the fourth card is a heart?

ANSWER

Since all four suits have the same number of cards, that means all four suits are equally likely to be the suit of the fourth card from the top. This means the probability the fourth card is a heart is 1/4.

$\frac{1}{4}$

16. If fourteen cards are dealt from the top of a well shuffled deck of cards one after another without replacement, what is the chance that the last card is a heart?

ANSWER

Since all four suits have the same number of cards, that means all four suits are equally likely to be the suit of the fourteenth card. This means the probability the fourteenth card is a heart is 1/4.

$\frac{1}{4}$

17. If fourteen cards are dealt from the top of a well shuffled deck of cards one after another without replacement, what is the chance that the second and fifth cards ares hearts?

ANSWER

Use the multiplication rule:

$$E(XI_N|K) = E(X|A\&K)P(A|K)$$

as it applies to the special case of probability:

$$P(A\&B|C) = P(A|B\&C)P(B|C).$$

Let A be the event or statement that the second card is a heart and let B be the event or statement that the fifth card is a heart. Take C to be the statement that all cards are equally likely to be anywhere in the deck unless we are given specific information otherwise. Then

$$P(A|C) = \frac{1}{4}$$

whereas

$$P(A|B\&C) = \frac{12}{51} = P(B|A\&C).$$

We therefore conclude by the multiplication rule that

$$P(A\&B|C) = P(A|B\&C)P(B|C) = (\frac{12}{51})(\frac{1}{4}) = \frac{3}{51}.$$

$\frac{3}{51}$

18. If fourteen cards are dealt from the top of a well shuffled deck of cards one after another without replacement, what is the chance that the fourth is a heart, given that the fifth is a spade?

ANSWER

$\frac{13}{51}$

19. If fourteen cards are dealt from the top of a well shuffled deck of cards one after another without replacement, what is the chance that the fourth is a heart given that the third is a spade?

ANSWER

$\frac{13}{51}$

20. If fourteen cards are dealt from the top of a well shuffled deck of cards one after another without replacement, what is the chance that the fourth is a heart given that the third is a heart, the fifth is a heart, the sixth is a club, and the seventh is a club?

ANSWER

$\frac{11}{48}$

21. A box contains 5 red blocks, 8 blue blocks, and 7 green blocks. A lab assisitant who is totally color blind is told to remove the blocks from the box one by one without replacement. What is the probability that the first block drawn is red?

ANSWER
$\frac{5}{20} = \frac{1}{4}$

22. A box contains 5 red blocks, 8 blue blocks, and 7 green blocks. A lab assisitant who is totally color blind is told to remove the blocks from the box one by one without replacement. What is the probability that the last block drawn is red?

ANSWER
$\frac{1}{4}$

23. A box contains 5 red blocks, 8 blue blocks, and 7 green blocks. A lab assisitant who is totally color blind is told to remove the blocks from the box one by one without replacement. What is the probability that the fourth block drawn is red?

ANSWER
$\frac{1}{4}$

24. A box contains 5 red blocks, 8 blue blocks, and 7 green blocks. A lab assisitant who is totally color blind is told to remove the blocks from the box one by one without replacement. What is the probability that the second block drawn is red given that the first is red?

ANSWER
$\frac{4}{19}$

25. A box contains 5 red blocks, 8 blue blocks, and 7 green blocks. A lab assisitant who is totally color blind is told to remove the blocks from the box one by one without replacement. What is the probability that the first block drawn is red given that the second is red?

ANSWER
$\frac{4}{19}$

26. A box contains 5 red blocks, 8 blue blocks, and 7 green blocks. A lab assisitant who is totally color blind is told to remove the blocks from the box one by one without replacement. What is the probability that the second block drawn is red given that the first and third are both red?

ANSWER
$\frac{3}{18}$

27. A box contains 5 red blocks, 8 blue blocks, and 7 green blocks. A lab assisitant who is totally color blind is told to remove the blocks from the box one by one without replacement. What is the probability that the fifth block drawn is red given that the first and third are both red and the seventh is green?

ANSWER
$\frac{3}{17}$

28. A box contains 5 red blocks, 8 blue blocks, and 7 green blocks. A lab assistant who is totally color blind is told to remove the blocks from the box one by one without replacement. What is the probability that the fifth block drawn is red given that the first and third are both red the seventh is green and the eighth is blue?

ANSWER
$\frac{3}{16}$

29. A box contains 5 red blocks, 8 blue blocks, and 7 green blocks. A lab assisitant who is totally color blind is told to remove the blocks from the box one by one without replacement. What is the probability that the fifth and ninth blocks drawn are both red given that the first and third are both red the seventh is green and the eighth is blue?

ANSWER

Let C be the statement of the problem setup and that the first and third are both red, the seventh is green and the eighth is blue. Let A be the statement that the ninth is red and let B be the statement that the fifth block drawn is red. From the multiplication rule of probability we know:

$$P(A\&B|C) = P(A|B\&C)P(B|C) = (\frac{2}{15})(\frac{3}{16}) = \frac{1}{40}.$$

$\frac{1}{40}$

30. Suppose that A, B and C are statements and that $P(A|B\&C) = .4$, that $P(A|C) = .7$, and that $P(B|C) = .6$. Calculate $P(B|A\&C)$.

ANSWER

We know:

$$P(A|B\&C)P(B|C) = P(A\&B|C) = P(B\&A|C) = P(B|A\&C)P(A|C),$$

and therefore putting in the numbers,

$$(.4)(.6) = (.7)P(B|A\&C),$$

so

$$P(B|A\&C) = \frac{.24}{.7} = \frac{24}{70} = \frac{12}{35}$$

$\frac{12}{35}$ or .3428571429

31. Suppose that A and B are statements and set $C =$ *not* B. Suppose $P(A|B) = .7$ and $P(A|C) = .8$. Finally, suppose that $P(B) = .3$. Calculate $P(A\&B)$.

ANSWER: Using the **Multiplication Rule**, $P(A\&B) = P(A|B)P(B) = (.7)(.3) = .21$.

FINAL ANSWER: .21

32. Suppose that A and B are statements and set $C =$ *not* B. Suppose $P(A|B) = .7$ and $P(A|C) = .8$. Finally, suppose that $P(B) = .3$. Calculate $P(A\&C)$.

ANSWER: Using $P(C) = 1 - P(B) = 1 - (.3) = .7$ and the **Multiplication Rule**,

$$P(A\&C) = P(A|C)P(C) = (.8)(.7) = .56.$$

FINAL ANSWER: .56

33. Suppose that A and B are statements and set $C =$ *not* B. Suppose $P(A|B) = .7$ and $P(A|C) = .8$. Finally, suppose that $P(B) = .3$. Calculate $P(A)$.

ANSWER: Using results from the previous problems

$$P(A) = P(A\&B) + P(A\&C) = .21 + .56 = .77.$$

FINAL ANSWER: .77

34. Suppose that A and B are statements and set $C =$ *not* B. Suppose $P(A|B) = .7$ and $P(A|C) = .8$. Finally, suppose that $P(B) = .3$. Calculate $P(B|A)$.

ANSWER: Using the **Multiplication Rule**, $P(B|A)P(A) = P(A\&B) = .21$, and $P(A) = .77$, so

$$P(B|A) = \frac{.21}{.77} = \frac{21}{77} = \frac{3}{11}.$$

FINAL ANSWER: $\frac{3}{11}$

35. Suppose A and B are statements with $P(A) = .4$ and $P(B) = .3$. Calculate $P(A \text{ or } B)$ assuming $P(A\&B) = .2$.

ANSWER: Using the **General Probability Rule for OR** we have

$$P(A \text{ or } B) = P(A) + P(B) - P(A\&B) = .4 + .3 - .2 = .5 = \frac{1}{2}.$$

FINAL ANSWER: $.5 = \frac{1}{2}$

36. Suppose A and B are statements with $P(A) = .4$ and $P(B) = .3$. Calculate $P(A \text{ or } B)$ assuming A and B are exclusive.

ANSWER: The assumption that A and B are exclusive tells us that $P(A\&B) = 0$ so $P(A \text{ or } B) = P(A) + P(B) = .4 + .3 = .7$

FINAL ANSWER: .7

37. Suppose A and B are statements with $P(A) = .4$ and $P(B) = .3$. Calculate $P(A \text{ or } B)$ assuming A and B are independent.

ANSWER: The assumption that A and B are independent tells us $P(A\&B) = P(A)P(B) = (.3)(.4) = .12$, so

$$P(A \text{ or } B) = .3 + .4 - .12 = .70 - .12 = .58.$$

FINAL ANSWER: .58

38. Suppose that a box contains 30 blocks. How many ways are there to take 5 blocks from the box?

ANSWER: There are

$$C(30, 5) = \frac{(30)(29)(28)(27)(26)}{(5)(4)(3)(2)(1)} = (6)(29)(7)(9)(13) = 142506 = (30\ nCr\ 5).$$

FINAL ANSWER: 142506

39. Suppose that a one box contains 20 blocks and another box contains 30 blocks. If the contents of both boxes are dumped into a third box, how many blocks will be in the third box?

ANSWER: Obviously $20 + 30 = 50$.

FINAL ANSWER: 50

40. Suppose that a box contains 70 RED blocks and ONE BLUE BLOCK for a total of 71 blocks. How many ways are there to draw 20 blocks from the box so as to have the blue block as one of the blocks drawn?

ANSWER: If you have to have the blue block, then you are really only drawing 19 blocks from the 70 red blocks, so the answer is 70 nCr 19 or about $(6.348)(10^{16})$.

FINAL ANSWER: $(6.348)(10^{16})$

41. Suppose that a box contains 70 RED blocks and ONE BLUE BLOCK for a total of 71 blocks. How many ways are there to draw 20 blocks from the box so as to NOT have the blue block as one of the blocks drawn?

ANSWER: If you have to NOT get the blue block, then all your 20 chosen blocks must come from the 70 red blocks so that would be 70 nCr 20 or about $(1.619)(10^{17}$.

FINAL ANSWER: $(1.619)(10^{17}$

42. Suppose that a box contains 70 RED blocks and ONE BLUE BLOCK for a total of 71 blocks. How many ways are there to draw 20 blocks from the box IF WE DON NOT CARE WHETHER OR NOT the blue block is one of the blocks drawn?

ANSWER: If you do not care whether or not you get the blue block there are 71 nCr 20 ways or about $(2.254)(10^{17}$. Notice this is approximately the sum of the previous two answers. In fact, the correct answer here is exactly the sum of the exact answers to the previous two questions, but the calculator cannot calculate these big numbers accurately.

FINAL ANSWER: $(2.254)(10^{17}$

43. Supose that we have a bucket with ten numbered blocks numbered one through ten and we have three empty numbered buckets numbered one two and three. How many ways are there to put 3 blocks in the first bucket, 2 blocks in the second bucket, and 2 blocks in the third bucket?

ANSWER: We do not care about the order within each bucket, and implicitly the left over blocks are in the starting bucket, so we are arranging ten numbered things into four buckets so as to have 3 in the first, 2 in the second, 2 in the third, and 3 in the last or starting bucket. If you were to do this and then order the blocks in each bucket, the result would be 10! ways. Thus, if x is the number of ways to put the blocks into the buckets, then

$$x(3!)(2!)(2!)(3!) = 10!$$

is the total nbumber of arrangements of the ten blocks, and therefore

$$x = \frac{10!}{(3!)(2!)(2!)(3!)} = 25200.$$

FINAL ANSWER: 50400

44. How many ways are there to make a 4 card hand which contains two pair but not four of a kind?

ANSWER: Obviously there are 13 four card hands which have four of a kind. To make a hand with two pair and not four of a kind, you must choose two denominations which can be done in $C(13, 2) = (6)(13) = 78$ ways, and then choose two cards of the first denomination which can be done in $C(4, 2) = 6$ ways and then choose two cards of the second chosen denomination which can be done again in six ways, so the total number of ways by the Multiplication Principle is therefore (78)(6)(6)=2808.

FINAL ANSWER: 2808

45. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. What is $Var(X)$, the variance of X?

ANSWER: The Variance is the square of the standard deviation $= \sigma$ so here we have the variance of X is $Var(X) = \sigma_X^2 = 7^2 = 49$.

FINAL ANSWER: 49

46. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. What is the mean of X^2?

ANSWER: In general we know that

$$E(X^2) = (\mu_X)^2 + (\sigma_X)^2$$

for any unknown X, so in our case at hand, we have

$$E(X^2) = (70)^2 + 7^2 = 4900 + 49 = 4949.$$

FINAL ANSWER: 4949

47. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. What is $Cov(X, Y)$, the covariance of X and Y?

ANSWER: In general we know that for any two unknowns we have

$$Cov(X, Y) = E([X - \mu_X][Y - \mu_Y]) = E(XY) - \mu_X\mu_Y = \rho \cdot \sigma_X \cdot \sigma_Y.$$

In our case, from the information we are given we would use $Cov(X, Y) = \rho\sigma_X\sigma_Y$ to find the covariance, so we have

$$Cov(X, Y) = \rho \cdot \sigma_X \cdot \sigma_Y = (.6) \cdot 7 \cdot 8 \cdot = 33.6$$

FINAL ANSWER: 33.6

48. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. What is $E(XY)$, the mean of XY?

ANSWER: In general we know that for any two unknowns we have

$$Cov(X, Y) = E([X - \mu_X][Y - \mu_Y]) = E(XY) - \mu_X\mu_Y = \rho \cdot \sigma_X \cdot \sigma_Y.$$

In our case, from the information we are given as we just found $Cov(X, Y) = \rho\sigma_X\sigma_Y = 33.6$, to find the mean of the product XY, we just use

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

and solve for $E(XY)$ finding

$$E(XY) = \mu_X \cdot \mu_Y + Cov(X, Y) = (70) \cdot (100) + 33.6 = 7000 + 33.6 = 7033.6.$$

FINAL ANSWER: 7033.6

49. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. What is the value of D_X, the deviation of X, if we are given that the value of X is 85?

ANSWER: By definition, the deviation of X, denoted D_X is given by $D_X = X - \mu_X$, so here we have

$$D_X = X - 70,$$

so if we are given the value of X is 84, then the value of D_X is $84 - 70 = 14$.

FINAL ANSWER: 15

50. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. What is $E(D_X^2)$, the expected squared deviation of X from its mean?

ANSWER: Remember, by definition,

$$Cov(X, Y) = E(D_X \cdot D_Y)$$

and therefore $\sigma_X^2 = Var(X) = Cov(X, X) = E(D_X^2)$, so here we have simply

$$E(D_X^2) = 7^2 = 49.$$

FINAL ANSWER: 49

51. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. What is the value of Z_X, the standardization of X, if we are given that the value of X is 84?

ANSWER: By definition, we have

$$Z_X = \frac{X - \mu_X}{\sigma_X} = \frac{D_X}{\sigma_X},$$

so if we are given the value of X is 84, the value of D_X is 14, so the value of Z_X is

$$\frac{84 - 70}{7} = \frac{14}{7} = 2.$$

FINAL ANSWER: 2

52. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. If we think the value of D_Y, deviation of Y, is 16, then what should we think is the value of Y?

ANSWER: Since $D_Y = Y - \mu_Y$, it follows that $Y = \mu_Y + D_Y$, so in the case at hand we have $Y = 100 + D_Y$, so if we think that D_Y is 16, then we should think Y is

$$100 + 16 = 116.$$

FINAL ANSWER: 116

53. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. What is $E(Y|D_Y = 16)$?

ANSWER: If we are given that $D_Y = 16$, then we know $Y = 100 + 16 = 116$, so it must be the case that

$$E(Y|D_Y = 16) = 116.$$

FINAL ANSWER: 116

54. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. If we think the value of Z_Y, the standardization of Y, is 2, then what should we think is the value of D_Y, the deviation of Y?

ANSWER: Since

$$Z_Y = \frac{D_Y}{\sigma_Y},$$

it follows that

$$D_Y = \sigma_Y \cdot Z_Y,$$

so here, as $\sigma_Y = 8$, we should use

$$D_Y = 8 \cdot Z_Y$$

and therefore if we think Z_Y, the standardization of Y, has value 2, then we should think D_Y, the deviation of Y, has value $8 \cdot 2 = 16$.

FINAL ANSWER: 16

55. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. If we think that Z_Y has value 2, then what should we think is the value of Y,?

ANSWER: Since $Y = \mu_Y + D_Y$ and since $D_Y = \sigma_Y \cdot Z_Y$, it follows that

$$Y = \mu_Y + \sigma_Y cdot Z_Y.$$

This means here that

$$Y = 100 + 8 \cdot Z_Y,$$

and therefore if we think that Z_Y has value 2, then we should think Y has value

$$100 + 8 \cdot 2 = 100 + 16 = 116.$$

FINAL ANSWER: 116

56. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. What is $E(Y|Z_Y = 2)$?

ANSWER: If you are given that $Z_Y = 2$, since here we know that $Y = 100 + 8 \cdot Z_Y$, we therefore have

$$E(Y|Z_Y = 2) = 100 + 8 \cdot 2 = 116.$$

FINAL ANSWER: 116

57. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. If we are given that $Z_X = 4$, then what should we guess for the value of Z_Y, that is, what is $E(Z_Y|Z_X = 4)$?

ANSWER: Here is where ρ, the correlation, comes in to play. In general,

$$E(Z_Y|Z_X = c) = \rho \cdot c.$$

That is, whatever we are given for the standard score for X we just multiply by ρ to get the optimal guess for the standard score for Y. If we are given that the value of Z_X is 4, then we should guess the standard score for Y is

$$E(Z_Y|Z_X = 4) = \rho \cdot 4 = (.6) \cdot 4 = 2.4.$$

FINAL ANSWER: 2.4

58. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. If we are told that the value of X is exactly the mean, which is 70, then what should we guess is the value of Y, that is, what is $E(Y|X = 70)$?

ANSWER: Notice that being equal to the true mean is the same as having standard score zero. Since zero multiplied by ρ is still zero, it follows that we should guess that Y is also exactly average, that is we should guess that Y has the value 100. That is

$$E(Y|X = 70) = 100.$$

FINAL ANSWER: 100

59. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. If we are given that $X = 98$, then what should we guess for the value of Y, that is, what is $E(Y|X = 98)$?

ANSWER: Since $X = 98$, that means $D_X = 98 - 70 = 28$, and therefore $Z_X = 28/7 = 4$, that is, we are therefore given $Z_X = 4$. This means we should guess Z_Y has value $.6 \cdot 4 = 2.4$. Thus, we should guess Y has value

$$100 + 8 \cdot 2.4 = 100 + 19.6 = 119.2.$$

Notice that in general, since

$$E(Z_Y|Z_X = z) = \rho \cdot z,$$

and since stating $X = x$ is the same as stating that $Z_X = z$, where

$$z = \frac{x - \mu_X}{\sigma_X},$$

it follows that

$$E(Y|X = x) = \mu_Y + \sigma_Y \cdot \rho \cdot \frac{x - \mu_X}{\sigma_X} = \mu_Y + \rho \cdot \frac{\sigma_Y}{\sigma_X}(x - \mu_X).$$

It is customary to define

$$\beta = \rho\frac{\sigma_Y}{\sigma_X}$$

and call β the **Regression Slope**, so we then have

$$E(Y|X = x) = \mu_Y + \beta \cdot (x - \mu_X)$$

for the optimal guess for Y given te value of X. Notice that if we set $y = E(Y|X = x)$, then we have simply

$$y = \mu_Y + \beta \cdot (x - \mu_X),$$

which is simply the equation of a straight line in the $(x, y)-$plane having slope β. To find its vertical intercept we just put in $x = 0$, which gives the value customarily denoted α given by

$$\alpha = \mu_Y - \beta \cdot \mu_X.$$

We call α the **Regression Intercept**. We therefore have the equivalent equations

$$y = \mu_Y + \beta(x - \mu_X)$$

and

$$y = \alpha + \beta x.$$

The first is obviously a line of slope β passing through the point (μ_X, μ_Y) whereas the second is called the slope-intercept form which is often faster for computing. Here, $\beta = (.6)(8/7) = 4.8/7 = 48/70$, and $\alpha = 100 - 70(48/70) = 100 - 480 = -380$, so the first equation is

$$y = 100 + \frac{48}{70}(x - 70)$$

and the second equation is

$$y = -380 + \frac{48}{70}x.$$

If you put $x = 98$ into either of these equations the resulting value for y is $y = 119.2$.

FINAL ANSWER: 119.2

60. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. If you guess that Y has value 119.2 when you know that $X = 98$, then what is your expected squared error?

ANSWER: If you use the information as to the value of X in this situation, using the regression equations to calculate the value of y to guess for Y, then your expected squared error is

$$E(R^2) = (1 - \rho^2) \cdot \sigma_Y^2.$$

Here, if $W = \alpha + \beta \cdot X$, then $R = W - Y$. That is, if W is what you use to guess Y from X using the regression equations, then $R = W - Y$ is your error. Thus in our case we have

$$E(R^2) = (1 - \rho^2) \cdot \sigma_Y^2 = (1 - [.6]^2) \cdot 8^2 = (.64) \cdot (64) = 64^2/100 = 4096/100 = 40.96$$

FINAL ANSWER: 40.96

61. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. If we guess the value of Y without using the value of X, then what is our expected squared error?

ANSWER: If we do not use the value of X, then we would guess that the value of Y is 100 and then our error is D_Y so our expected squared error is $E(D_Y^2) = \sigma_Y^2 = 64$. Notice that 40.96 is a lot less than 64. Using knowledge of X to aid in guessing the value of Y is substantially cutting the error down when $\rho = .6$. Thus, in general, ρ^2 is the fraction of variation in Y that is accounted for with the regression equation and thus $1 - \rho^2$ is the fraction of variation in Y that is left not accounted for with the regression equation.

FINAL ANSWER: 64

62. Suppose that X and Y are unknowns and that we know the regression slope is $\beta = 5$ for the regression line equation. Suppose that we also know that if X has value 60 then the regression equation tells us to guess 90 for the value of Y. If we are given that $X = 62$, then what should we guess for Y?

ANSWER: Since the regression slope $\beta = 5$, this means that for each unit increase in X we should increase our guess for Y by 5 units. Thus as we know 62 is a 2 unit increase for X we should increase our guess for Y by $5 \cdot 2 = 10$ units from 90 up to 100.

FINAL ANSWER: 100

63. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. What is $Var(X + Y)$, the variance of $X + Y$?

ANSWER: In general, as

$$\sigma_X^2 = Var(X) = Cov(X, X),$$

you should think of covariance as like the multiplication for which variance is squaring. The algebraic formula

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

corresponds to the formula

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

or

$$Var(X \pm Y) = \sigma_X^2 + \sigma_Y^2 \pm 2 \cdot \rho \cdot \sigma_X \cdot \sigma_Y.$$

Here we then have

$$Var(X + Y) = 7^2 + 8^2 + 2 \cdot (.6) \cdot 7 \cdot 8 = 49 + 64 + 2 \cdot 33.6 = 49 + 64 + 67.2 = 180.2.$$

FINAL ANSWER: 180.2

64. Suppose that X and Y are unknowns and that $E(X) = \mu_X = 70$, that $E(Y) = \mu_Y = 100$, that $\sigma_X = 7$, that $\sigma_Y = 8$, and that the correlation between X and Y is $\rho = .6$. What is $Var(X - Y)$, the variance of $X - Y$?

ANSWER: In general, as

$$\sigma_X^2 = Var(X) = Cov(X, X),$$

you should think of covariance as like the multiplication for which variance is squaring. The algebraic formula

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

corresponds to the formula

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

or

$$Var(X \pm Y) = \sigma_X^2 + \sigma_Y^2 \pm 2 \cdot \rho \cdot \sigma_X \cdot \sigma_Y.$$

Here we then have

$$Var(X - Y) = 7^2 + 8^2 - 2 \cdot (.6) \cdot 7 \cdot 8 = 49 - 64 - 2 \cdot 33.6 = 49 + 64 - 67.2 = 45.8.$$

Notice that it is only the last term which is affected by the minus sign.

FINAL ANSWER: 45.8

65. Suppose that a pond contains 50 fish of which 30 are redfish. If we catch 10 without replacement forming a simple random sample of fish in the pond, what is the probability that the number of redfish we find in our sample is exactly 7?

ANSWER: Since we are sampling without replacement, if T denotes the number of redfish in our sample, then T has the **Hypergeometric Distribution** with population size $N = 50$, sample size $n = 10$, and population of success size $R = 30$. Therefore

$$P(T = k) = (R\ nCr\ k)([N - R]\ nCr\ [n - k])/(N\ nCr\ n),$$

and here this is

$$P(T = 7) = (30\ nCr\ 7)(20\ nCr\ 3)/(50\ nCr\ 10) = .2259296294,$$

or about .226, to three significant digits.

FINAL ANSWER: .226

66. Suppose that a pond contains 50 fish of which 30 are redfish. If we catch 10 **with** replacement, throwing the fish back after each catch so as to form an independent random sample, then what is the probability that exactly 7 of the 10 fish caught will be redfish?

ANSWER: Since the sampling is independent random sampling, the sucessive catches are independent and the total number T of redfish caught obeys the **Binomial Distribution** with sample size $n = 10$ and success rate $p = 30/50 = .6$. Thus now

$$P(T = k) = (n\ nCrk)p^k(1 - p)^{n-k} = binompdf(n, p, k)$$

and here this is

$$P(T = 7) = binompdf(10, 30/50, 7) = .214990848,$$

or about .215, to three significant digits.

FINAL ANSWER: .215

67. Suppose that a pond contains 50 fish of which 30 are redfish. If we catch 10 **with** replacement, throwing the fish back after each catch so as to form an independent random sample, then what is the probability that no more than 7 of the 10 fish caught will be redfish?

ANSWER: Here again we have the total number T of redfish caught obeys the binomial distribution, but now we need to calculate

$$P(T \leq 7) =$$
$$P(T = 0) + P(T = 1) + P(T = 2) + ... + P(T = 6) + P(T = 7)$$
$$= binomcdf(10, .6, 7) = .8327102464,$$

or about .837, to three significant digits.

FINAL ANSWER: .837

68. Suppose that our pond on average contains 20 tadpoles per cubic foot of water. We have a bucket which holds 2.5 cubic feet of water. Assume that the number of tadpoles in disjoint regions of water in the pond are independent. How many tadpoles do we expect to find in a bucket full of pond water from our pond?

ANSWER: Since on average there are 20 tadpoles per cubic foot and since our bucket holds 2.5 cubic feet of pond water, we expect that when filled with water from our pond it will contain $\mu = (20) \cdot (2.5) = 50$ tadpoles.

FINAL ANSWER: 50

69. Suppose that our pond on average contains 20 tadpoles per cubic foot of water. We have a bucket which holds 2.5 cubic feet of water. Assume that the number of tadpoles in disjoint regions of water in the pond are independent. What is the probability that a bucket full of our pond water contains exactly 47 tadpoles?

ANSWER: If T denotes the number of tadpoles we will find, then $\mu_T = (20) \cdot (2.5) = 50$, and T obeys the **Poisson Distribution**, so

$$P(T = k) = \frac{e^{-\mu_T}(\mu_T)^k}{k!} = poissonpdf(\mu_T, k).$$

Here then

$$P(T = 47) = \frac{e^{-50}(50)^{47}}{47!} = poissonpdf(50, 47) = .052990566,$$

or about .0530, to three significant digits.

FINAL ANSWER: .0530

70. Suppose that our pond on average contains 20 tadpoles per cubic foot of water. We have a bucket which holds 2.5 cubic feet of water. Assume that the number of tadpoles in disjoint regions of water in the pond are independent. What is the probability that a bucket full of our pond water no more than 47 tadpoles?

ANSWER: Again, T, the number of tadpoles in our bucket of pond water obeys the Poisson distribution, so

$$P(T \leq 47) =$$
$$P(T = 0) + P(T = 1) + P(T = 2) + ... + P(T = 46) + P(T = 47)$$
$$= poissoncdf((20) \cdot (2.5), 47) = poissoncdf(50, 47) = .3696681718,$$

or about .370, to three significant digits.

FINAL ANSWER: .370

71. Suppose that the fish in my pond have mean length 20 inches with a standard deviation of 4 inches, and that we know nothing else about the distribution of lengths for this population of fish. What is the probability that a fish caught from this pond will be over 20 inches in length?

ANSWER: Since we only know the mean and standard deviation and nothing else about the distribution, the distribution we must use is the normal distribution. Use of any other distribution would be assuming more knowledge than we have. But, any normal distribution is symmetric about its mean, that is, if X is normal, then

$$P(X \geq \mu_X) = P(X > \mu_X) = \frac{1}{2},$$

so as we are given here $\mu_X = 20$ and asked for $P(X > 20)$, the answer is simply one half.

FINAL ANSWER: $\frac{1}{2}$

72. Suppose that the fish in my pond have mean length 20 inches with a standard deviation of 4 inches, and that we know nothing else about the distribution of lengths for this population of fish. What is the probability that a fish caught from this pond will be between 22 and 27 inches in length?

ANSWER: Since we only know the mean and standard deviation and nothing else about the distribution, the distribution we must use is the normal distribution. Use of any other distribution would be assuming more knowledge than we have. We can use the normal distribution in the calculator, to calculate the area under the normal curve with mean 20 and standard deviation between the limits 22 and 27. Using X to denote fish length, we have

$$P(22 \leq X \leq 27) = P(22 < X < 27) = \frac{1}{4 \cdot \sqrt{2\pi}} \int_{22}^{27} e^{-\frac{1}{2}\left(\frac{x}{4}\right)^2} dx$$
$$= \mathit{normalcdf}(22, 27, 20, 4) = .2684784187,$$

or about .268, to three significant digits.

FINAL ANSWER: .268

73. Suppose that the fish in my pond have mean length 20 inches with a standard deviation of 4 inches, and that we know nothing else about the distribution of lengths for this population of fish. What is the probability that an independent random sample of 5 fish caught from this pond will average between 19 and 23 inches in length?

ANSWER: Since we only know the mean and standard deviation and nothing else about the distribution, the distribution we must use is the normal distribution. Use of any other distribution would be assuming more knowledge than we have. We know that the sample mean $\bar{X}$ is expected to be $\mu_X = 20$, and for independent random sampling, the standard deviation of $\bar{X}_n$ is

$$SD(\bar{X}_n) = \frac{\sigma_X}{\sqrt{n}},$$

so here $\bar{X}$ is normal with mean 20 and standard deviation $4/\sqrt{5}$. Thus, here

$$P(19 \leq \bar{X} \leq 23) = \mathit{normalcdf}(19, 23, 20, 4/\sqrt{(5)}) = .6651587438,$$

or about .665, to three significant digits.

FINAL ANSWER: .665

74 Suppose that the fish in my pond have mean length 20 inches with a standard deviation of 4 inches, and that we know nothing else about the distribution of lengths for this population of fish. What is the probability that in an independent random sample, the total length of 5 fish caught from this pond will be between 95 and 110 inches in length?

ANSWER: Since we only know the mean and standard deviation and nothing else about the distribution, the distribution we must use is the normal distribution. Use of any other distribution would be assuming more knowledge than we have. For a sample of size n we have the expected value of T_n, the sample total, is

$$E(T_n) = n \cdot \mu_X,$$

so here we have

$$E(T_5) = 5 \cdot (20) = 100,$$

and for independent random sampling, we have the standard deviation $SD(T_n)$ of T_n is

$$SD(T_n) = \sqrt{n} \cdot \sigma_X,$$

which here is

$$SD(T_5) = \sqrt{5} \cdot (4).$$

Moreover, as X here is normal, the same is true for T_n and $\bar{X}$, so

$$P(95 \leq T_n \leq 110) = \mathit{normalcdf}(95, 110, 100, 4 \cdot \sqrt{(5)}\ ) = .5801486913,$$

or about .580, to three significant digits.

FINAL ANSWER: .580

75. Suppose that buses arrive at my bus stop at an average rate of 3 per hour, which is the same as one every twenty minutes on average. Let W be the time I have to wait at my bus stop for a bus. What is the probability that I must wait more than 25 minutes for a bus? Assume that the number of buses arriving during disjoint time intervals are independent.

ANSWER: The independence assumption means that the number of buses arriving during any interval of time is governed by the Poisson distribution. Thus, if T is an interval of time in minutes, and if X is the number of buses that arrive during time T, then $E(X) = T/(20)$. Thus, in 25 minutes I expect to see $25/20 = 1.25$ buses, but if I have to wait more than 25 minutes, it is because during that 25 minutes I actually saw no buses, and therefore

$$P(W > 25) = \mathit{poissoncdf}(25/20, 0) = \mathit{poisssoncdf}(1.25, 0) = .2865047969,$$

or about .287, to three significant digits. Notice since we are calculating the probability of seeing zero buses, it does not matter whether we use the cdf or the pdf, the answer is the same, as the number of buses we see cannot be negative.

FINAL ANSWER: .287

76. Suppose that the fish in my pond have mean length 20 inches with a standard deviation of 4 inches, and that we know nothing else about the distribution of lengths for this population of fish. What is the shortest a fish can possibly be and still be in the top 20 percent of fish as regards length?

ANSWER: Since we only know the mean and standard deviation and nothing else about the distribution, the distribution we must use is the normal distribution. Use of any other distribution would be assuming more knowledge than we have. Now, we are not asking for a probability here, we are asking the reverse, namely you must find the value of x so that $P(X \leq x) = .8$ and $P(X \geq x) = .2$. It is customary to denote this x as $x_{.8}$, and to call this the 80^{th} percentile score. We are in the situation that we have the probability, and we want the limit. This in a sense is going backwards and that is the job of "invNorm" in your distribution menu. Thus here,

$$x_{.8} = invNorm(.8, 20, 4) = 23.36648493,$$

or about 23.4 inches long, to three significant digits. Notice that there may in fact be only one fish in the pond and these answers are the same, that is we are dealing with the state of our information. We can say that if a fish is caught there will be an 80 percent chance that it is less than 23.4 inches long.

FINAL ANSWER: 23.4

77. Suppose that I do not know the mean weight of the fish in my pond and I want to estimate it using the sample mean of an independent random sample. Suppose that I have a sample of 36 fish with sample mean $\bar{x} = 8.3$ pounds. Suppose that I know that the standard deviation in fish weight for the fish in my pond is 2 pounds. What is the margin of error in the 95 percent confidence interval for the true mean weight of the fish in my pond?

ANSWER: Since the sample size is over 30 and the sample is an IRS, it follows that $\bar{X}$ is normally distributed with standard deviation $2/\sqrt{36} = 2/6 = 1/3$, and this means the margin of error in the 95 percent confidence interval is M where

$$M = z_{(C=.95)} \cdot \frac{\sigma_X}{\sqrt{n}} = z_{(C=.95)} \cdot \frac{1}{3}.$$

To find z_C, here we can use the invNorm in the distribution menu,

$$z_{(C=.95)} = invNorm(.975, 0, 1) = -invNorm(.025, 0, 1) = 1.959963986.$$

This means

$$M = .6533213287, \text{ or } .653,$$

to three significant digits.

Alternately, you can use the z-interval in the calculator. Choose the stats option enter the information as directed, but use $\bar{x} = 0$, choose the C-level as .95, and in the readout, the positive number is M, which in this case is $M = .65332$ which gives again $M = .653$ to three significant digits.

FINAL ANSWER: .653

78. Suppose that I do not know the mean weight of the fish in my pond and I want to estimate it using the sample mean of an independent random sample. Suppose that I have a sample of 36 fish with sample mean $\bar{x} = 8.3$ pounds. Suppose that I do not know the standard deviation in the weight of fish in my pond, but the sample standard deviation of my sample is 2 pounds. What is the margin of error in the 95 percent confidence interval for the true mean weight of the fish in my pond?

ANSWER: This problem is the same as the preceding except that the standard deviation of the population is unknown and we use the sample standard deviation in its place. Whenever you do that, you must use the Student-t-distribution for $n - 1 = 35$ degrees of freedom in place of the standard normal distribution. Thus, the margin of error is now

$$M = t_{(C=.95)} \cdot \frac{s}{\sqrt{n}} = t_{(C=.95)} \cdot \frac{1}{3}.$$

To find t_C we can use the invT in the calculator,

$$t_{(C=.95)} = invT(.975, 35) = 2.030107868,$$

so

$$M = .6767026228, \text{ or } .677,$$

to three significant digits. Notice that the margin of error is increased due to the lack of knowledge of the population standard deviation, even though the sample standard deviation accidentally hits the population standard deviation on the nose. But we do not know that here. Alternately, you can use the t-interval in the calculator. Choose the stats option and enter $\bar{x} = 0$ and $s_x = 2$, and $n = 36$, and then the readout is $(-.6767, .6767)$ which means $M = .6767$ or again .677 to three significant digits.

FINAL ANSWER: .677

79. Suppose that I do not know the percentage of ducks that will vote for Donald for Mayor of Duckburg in an upcoming election. Suppose that in a simple random sample of 2000 citizens asked, 1078 say they will vote for Donald in the election. What is the 95 percent confidence interval for the true proportion of citzens of Duckburg who say they will vote for Donald in the upcoming election?

ANSWER: Here, the sample data is used to calculate first the sample total number of yes votes for Donald which we can call T_n, and then the true proportion is estimated by the sample proportion usually denoted $\hat{p}$ given by

$$\hat{p} = \frac{1}{n} T_n = \bar{X}_n,$$

that is to say, the same thing is being done here as in the previous two problems, except that now the basic unknown we are dealing with is simply the unkown

$$X = I_D,$$

where D denotes the statement that the person asked answers yes to voting for Donald.

Now, using $p = P(D)$, for the probability of D, which is the true proportion of voters who support Donald, and since $X^2 = X$ as X is an indicator, and

$$E(X) = E(I_D) = P(D) = p,$$

it follows that

$$Var(X) = Var(I_D) = E(X^2) - [E(X)]^2 = E(X) - p^2 = p - p^2 = p(1-p) = P(D) \cdot P(not\ D).$$

Also, since

$$0 \leq p \leq 1,$$

it follows that

$$p(1-p) \leq (1/2)^2 \leq 1/4,$$

and therfore

$$\sigma_X \leq \frac{1}{2}.$$

Even though T_n is discreet here, with such large $n = 2000$, we can assume that T_n and $\hat{p} = \bar{X}$ are normal. Therefore, M, the margin of error with 95 percent confidence is definitely no more than

$$M \leq z_{(C=.95)} \cdot \frac{(1/2)}{\sqrt{n}} = (1.959963986) \frac{.5}{\sqrt{2000}} = .0219130635, \text{ or } .0219,$$

to three significant digits.

Now, for such a large sample, we can be fairly sure to be reaonalby close to the true proportion (we see we are 95 percent sure we are within .0219), and the standard deviation as a function of the true proportion p does not change very much for slight change in p, so it is accepted statistical practice to use the estimate $\hat{p}$ from the sample data to calculate the value of σ_X thus getting an estimate which we can denote $\hat{\sigma}$ for σ_X here, that is

$$\hat{\sigma} = \sqrt{\hat{p}(1-\hat{p})}.$$

This gives

$$M = z_{(C=.95)} \cdot \frac{\hat{\sigma}}{\sqrt{n}} = z_{(C=.95)} \cdot \sqrt{\frac{T_n(n-T_n)}{n^3}} = .0218463023, \text{ or } .0218.$$

Now, this last calculation is what is done in the calculator when you use the 1-Prop z-interval in the calculator's TEST menu. You enter the information and the readout gives

$$(.51715,\ .56085)$$
$$\hat{p} = .539$$
$$n = 2000.$$

This tells us that with 95 percent confidence

$$.5715 \leq p \leq .56085.$$

To find the margin of error here to compare this result with our previous result, notice that if the confidence interval for μ has the form (a, b), then that is the same as saying with confidence C, we know

$$a \leq \mu \leq b,$$

and if $\bar{x}$ is the value of the sample mean from the data, then

$$a = \bar{x} - M,$$

and

$$b = \bar{x} + M.$$

This means that if we know the confidence interval and we know the value of $\bar{x}$, then we easily find the margin of error

$$M = b - \bar{x}.$$

Here this means

$$M = .56085 - .539 = .02185,$$

which is again .219 to three significant digits.

However, this problem asked for the confidence interval itself, not the margin of error. Thus, we can get the answer from the readout directly, or calculate using

$$a = \hat{p} - M,\ \ b = \hat{p} + M,$$

and that result to three significant digits is agian

$$.517 \leq p \leq .561.$$

Returning to the margin of error here denoted M, we see that $M = .0219$ means that roughly we have an error of $2/100$, and we do not call this a two percent error, but rather an error of 2 percentage points. You should notice the distinction here. Whenever you speak of a percentage error you mean the fraction error divided by true amount, which would be rougly $.02/.539$ or about 4 percent here. That is here, an error of two percentage points is about a four percent error. However, when speaking of a percent error, we need to know the actual true value to know the percentage error, and we do not know that here, we can only guess its precise value to be around 4 percent. On the other hand, the margin of error for the confidence interval is very precisely known.

We should notice here that in using the normal approximation to the binomial, we should also use the continuity correction which we have here ignored and which your calculator ignores too. In fact, we should recall that if T_n is binomially distributed, then to approximate $P(a \leq T_n \leq b)$, where a and b are integers using the normal distribution, we use

$$P(a \leq T_n \leq b) = normalcdf(a - .5, b + .5, np, \sqrt{np(1-p)}\ ),$$

so we should really set this equal to the level of confidence C, so

$$C = normalcdf(a - .5, b + .5, np, \sqrt{np(1-p)}\ ).$$

Now, the margin of error in T_n is nM, and this means that the usual endpoints should be corrected to be

$$a - .5 = T_n - nM$$

and

$$b + .5 = T_n + nM,$$

and therefore the confidence interval for the proportion should really better be calculated as

$$a = \frac{T_n - nM + .5}{n} = \hat{p} - M + \frac{.5}{n}$$

and

$$b = \frac{T_n + nM - .5}{n} = \hat{p} + M - \frac{.5}{n}.$$

Now .5 divided by 2000 is .00025, so this slightly changes the confidence interval, but not in the significant digits. Thus, technically, our confidence interval is

$$(.539 - .0219130635 + .00025,\ .539 + .0219130635 - .00025) = (.5173369365,\ .5606630635)$$

which is the same result, to three significant digits. This means that for samples in the thousands, the confidence interval calculation can safely ignore the continuity correction in the normal approximation to the binomial.

FINAL ANSWER: $.517 \leq p \leq .561$

80. Suppose that I do not know the mean weight of the fish in my pond and I want to estimate it using the sample mean of an independent random sample. Suppose that I have a sample of 36 fish with sample mean $\bar{x} = 8.3$ pounds. Suppose that I know that the standard deviation in fish weight for the fish in my pond is 2 pounds. What is the significance of this data as evidence that the true mean weight of the fish in my pond actually exceeds 7.5 pounds?

ANSWER: As soon as we are asked for significance of data, we know we are dealing with a HYPOTHESIS TEST. Here we know the population standard deviations, so the significance is the value of p in the readout for the z-test in the test menu. We must choose the correct alternate hypothesis. Always remember, the alternate hypothesis is what you are trying to prove with the data. In this case, we are asking for the significance ($= P-$Value) as evidence that the true mean weight exceeds 7.5, so we enter $\mu_0 = 7.5$, and choose $\mu > \mu_0$ as our alternate hypothesis. The result is

$$P - Value = P(\bar{X}_n \geq \bar{x}_{data}) = P(\bar{X}_n \geq 8.3) = P(Z \geq z_{data}),$$

where z_{data} is the standard score for our sample mean under the null hypothesis H_0 which tells us to assume the true mean is $\mu_0 = 7.5$, so

$$H_0 : \mu_X \leq 7.5,$$

$$H_{alt} : \mu > 7.5,$$

$$z_{data} = \frac{8.3 - 7.5}{(2/\sqrt{36})} = \frac{6(.8)}{2} = 2.4,$$

$$P - Value = P(Z \geq 2.4) = .0081975289, \text{ or } .00820,$$

to three significant digits. Using the calculator test menu you begin by choosing the z-test, since you know the population standard deviation here. You then enter $\mu_0 = 7.5$, and $\sigma_x = 2$ as well as the data information asked for. Be sure to choose the correct alternate hypothesis in terms of μ_0. Thus, with $\mu_0 = 7.5$, the alternate hypothesis $\mu > 7.5$ becomes $\mu > \mu_0$, so we put the cursor on $> \mu_0$ and hit the enter button on the calculator. When we choose calculate and hit enter we see in the readout

$$Z - Test$$
$$\mu > 7.5$$
$$z = 2.4$$
$$p = .0081975289$$
$$\bar{x} = 8.3$$
$$n = 36$$

which again gives the same result as we calculated above for the significance of the data, again to three significant digits, the P-Value or significance of the data is .00820.

FINAL ANSWER: .00820

81. Suppose that I do not know the mean weight of the fish in my pond and I want to estimate it using the sample mean of an independent random sample. Suppose that I have a sample of 36 fish with sample mean $\bar{x} = 8.3$ pounds. Suppose that I do not know the standard deviation in the weight of fish in my pond, but the sample standard deviation of my sample is 2 pounds. What is the significance of this data as evidence that the true mean weight of the fish in my pond actually exceeds 7.5 pounds?

ANSWER: Notice that this is again a hypothesis test as we are asked for the significance ($= P-$Value) of the data. Just as for confidence intervals, when dealing with something you measure we must use z or t to test hypotheses concerning the mean, and the distinction between z and t is simply whether or not we know the true population standard deviation. Here we do not, and consequently we will use the $t-$ distribution for $n - 1 = 35$ degrees of freedom. Thus, we have

$$t_{data} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{8.3 - 7.5}{2/6} = 3 \cdot (.8) = 2.4$$

and

$$P{-}Value = P(t_{35} \geq 2.4) = .5 - P(0 \leq t_{35} \leq 2.4) = .5 - tcdf(0, 2.4, 35) = .0109249267, \text{ or } .0109,$$

to three significant digits. Using the t-test in the test menu, enter $\mu_0 = 7.5$ which is the value of the true mean μ_X we assume under the null hypothesis, and after entering the information asked for from the data stats, just choose again the correct alternate hypothesis just as for the z-test, and the readout gives

$$\begin{gathered} T - Test \\ \mu > 7.5 \\ t = 2.4 \\ p = .01090249267 \\ \bar{x} = 8.3 \\ s_x = 2 \\ n = 36 \end{gathered}$$

of which we see the value of p is what we really want as that is the significance or $P-$Value of the data.

Notice here that the significance of the data is numerically a larger number than in the previous problem even though all the numbers were the same. This is because in this problem we have less knowledge as we are using the sample standard deviation in place of the true population standard deviation which is unknown to us. Therefore our whole argument is weakened and our data becomes less significant, which means numerically a larger value for the P-Value. Keep this in mind: very significant data has a very small P-Value. When we speak of "highly significant evidence", we mean evidence or data with a very small P-Value, whereas data with a large P-Value, say above .2 is insignificant. Generally, a P-Value should be down at or below .05 before we pay attention to the data.

FINAL ANSWER: .0109

82. Suppose that I do not know the percentage of ducks that will vote for Donald for Mayor of Duckburg in an upcoming election. Suppose that in a simple random sample of 2000 citizens asked, 1078 say they will vote for Donald in the election. What is the significance of this data as evidence that the true proportion of citzens of Duckburg who say they will vote for Donald in the upcoming election actually exceeds .51?

ANSWER: Again, as we are asked for the significance of the data, we know we are dealing with a hypothesis test. In this case we are dealing with a question about the true proportion of citizens who say they will vote for Donald for Mayor of Duckburg. This means that our sample data is really consisting of the number of successes in n independent trials, where we regard it as a success when a citizen says he will vote for Donald for Mayor of Duckburg. The unknown here is the indicator of the event D, denoted I_D, just as in the previous problem where the estimate of the true proportion who say they will vote for Donald for Mayor of duckburg was made as a confidence interval with 95 percent confidence. Thus, the sample total T_n is binomially distributed. There, we used the fact that the binomial for large n is approximately normal. Likewise, we could do that here and in fact, that is what the 1-PropZ-test does, but it is not as accurate as it should be because the continuity correction for passing from the discrete count to the continuous normal has been left out, so the sample sizes have to be in the thousands for good accuracy.

To use the binomial distribution here is actually quite simple. Let p_D be the true proportion of citizens who say they will vote for Donald. Our hypothesis test is then

$$H_{ALT} : p_D > .51$$
$$H_0 : p_D \leq .51$$

and as usual, the inequality we calculate the probability of just mimics the alternate hypothesis. If T_n denotes the yes count, then it is binomially distributed and under the null hypothesis we assume the true proportion is $p_0 = .51$, so

$$P - Value = P(T_n \geq 1078 | T_n \; binomial, \; p_D = p_0, \; n = 2000)$$
$$= 1 - binomcdf(2000, .51, 1077) = .0050347322, \text{ or } .0503,$$

to three significant digits. If you use the 1-PropZTest in the test menu, the readout gives

$$1 - PropZTest$$
$$prop > .51$$
$$z = 2.594357777$$
$$p = .0047384234$$
$$\hat{p} = .539$$
$$n = 2000$$

and we see that the P-Value is off by quite a bit, since the answer to three significant digits is .00474, whereas with a direct calculation using the binomial distribution the result is .503 to three significant digits. To see what happened here, we have to ask what the calculator is actually calculating. It is trying to calculate the previous probability using the normal approximation to the binomial. The proper way to do that is to calculate

$$P - Value = P(T_n \geq 1078) = normalcdf(1077.5, \infty, (2000) \cdot (.51), \sqrt{(2000)(.51)(.49)})$$
$$= .5 - normalcdf((2000)(.51), 1077.5, (2000)(.51), \sqrt{(2000)(.51)(.49)})$$
$$= .0050557815, \text{ or } .506,$$

to three significant digits, which we see is only off in the third decimal place. If we replace the 1077.5 demanded by the continuity correction with 1078, in the previous calculation, the result is .0047384239, which agrees with the 1-PropZTest result to nine decimal places, so from that we can conclude that the calculator is programmed to ignore the continuity correction in using its normal approximation to the binomial for the z-test. This will cause error unless the sample size is around twenty thousand. On the other hand, the calculator's built in binomial distribution may also have error in it for large n, and to be able to choose between the .00503 and the .00506 would be difficult without knowing the details of the calculators program. These kinds of considerations must often be made in real life statistics, and the standard operating procedure should be to err on the conservative side. Here, that would mean choosing the value .00506 as that is larger and we do not want to make an unwarranted claim of a smaller P-Value. You might also notice here, that as far as the seriousness of an election is concerned, as the P-Value is very low in any one of these results, we would be inclined to think of the alternate hypothesis as fairly well established by this data. You always have to keep in mind that in many applications, the level of precision we are calculating here is often a waste of time in view of the overall implications of what you are doing. For instance, in dealing with elections, the significance of .005 is very strong evidence for the alternate hypothesis we are trying to prove, and there is virtually no real distinction in practical terms between the numbers .00474, .00503, and .00506. However, for your classwork in this class, we will stick to three significant digits in order to have a simple a uniform standard.

FINAL ANSWER: .00474, or .00503, or .00506

83. Suppose that I do not know the percentage of mice who support Goofy for Mayor of Duckburg but I have asked 10 mice and 7 say they support Goofy for Mayor of Duckburg whereas three support other candidates. What is the significance of this as evidence that the true proportion of mice who support Goofy for Mayor of Duckburg exceeds sixty percent?

ANSWER: This is a hypothesis test but the test statistic is the total count T of yes votes in an independent random sample of size only 10, so the count T is binomially distributed but the sample size is too small to apply the normal approximation to the binomial., that is, T may not be approximately normally distributed. This means that the binomial distribution in you calculator must be used to calculate the significance of this data, the 1-PropZTest in the TEST menu **CANNOT BE USED HERE.**

The hypothesis we are asking the data to prove is $H_A : p_G > .6$ so the null hypothesis is $H_0 : p_G \leq .6$, where p_G is the true percentage of mice that support Goofy for Mayor of Duckburg.

Our hypothesis test is:

$$H_0 : p_G \leq .6$$
$$\text{vs}$$
$$H_A : p_G > .6$$

and H_A of course is the alternate hypothesis whereas H_0 is the null hypothesis. Thus our null hypothesis tells us to assume the null hypothesis value of $p_0 = .6$ for the true proportion, so T is binomially distributed with $n = 10$ and success rate $p_0 = .6$. When we look at the alternate hypothesis, we see that the data significance is

$$\begin{aligned} P - Value = P(T \geq 7|n = 10, p_0 = .6) &= 1 - P(T < 7|n = 10, p_0 = .6) \\ &= 1 - P(T \leq 6|n = 10, p_0 = .6) \\ &= 1 - binomcdf(10, .6, 6) \\ &= .3822806016, \text{ or } .382, \end{aligned}$$

to three significant digits (and do not confuse significant digits with significant data). You can also notice here that this can be possibly more easily worked from the standpoint of mice who are against Goofy, as if p_A denotes the true proportion of mice who do not support Goofy for Mayor of Duckburg, then to say $p_G > .6$ is equivalent to saying $p_A < .4$, and our data tells us that 3 out of 10 are against Goofy. So the hypothesis test is also equivalent to the hypothesis test

$$H_0 : p_A \geq .4$$
$$\text{vs}$$
$$H_A : p_A < .4$$

and for this we assume the null hypothesis value of $p_0 = .4$ and looking at H_A now tells us to calculate the significance or P-Value as

$$P - Value = P(T \leq 3|n = 10, p_0 = .4) = binomcdf(10, .4, 3) = .3822806016, \text{ or } .382,$$

to three significant digits, the exact same result as before. Thus, you can always rearrange the hypothesis test logically so that the P-Value is calculated directly with the binomcdf in the calculator.

FINAL ANSWER: .382

84. Suppose that I want to make a 95 percent confidence interval for the true mean length of fish in my pond and I know that the length of fish in my pond is normally distributed with a standard deviation of at most $B = 8$ inches. What is the minimum size of an independent random sample from the population of fish in my pond required in order that the margin of error in my confidence interval be at most .25 inches?

ANSWER: Since the margin of error is M, where

$$M = \frac{z\sigma}{\sqrt{n}},\ z = invNorm(.975, 0, 1),$$

it follows that replacing σ by B will only increase M, so if we can choose n so as to make M be at most .25 using $B = 8$ in place of σ, then certainly the true σ will only serve to possibly reduce the margin of error but not increase it. Thus, in the equation

$$M = \frac{zB}{\sqrt{n}},$$

we simply solve for n and round up to the nearest whole number. The result is

$$n \geq \left(\frac{zB}{M}\right)^2,\ z = invNorm(.975, 0, 1),\ B = 8,\ M = .25,$$

so

$$n \geq \left(\frac{z \cdot 8}{.25}\right)^2 = 3933.653838,$$

which rounds up to 3934. Thus,

$$n \geq 3934.$$

FINAL ANSWER: 3934

85. Suppose that I want to make a 95 percent confidence interval for the true proportion of fish in my pond that are redfish. I want the margin of error to be at most .02. What is the miminum size of an independent random sample that will accomplish this.

ANSWER: The sampling for determination of a true proportion p is a special case of sampling a random variable which is an indicator of an event A, in this case, the event is that a randomly chosen fish in my pond is a redfish, and

$$\mu = E(I_A) = P(A) = p.$$

The sample mean in this case is the sample proportion, denoted $\hat{p}$, so if $X = I_A$ is the indicator, then T_n is the count of redfish in the sample and

$$\bar{X}_n = \frac{T_n}{n} = \hat{p}$$

is the sample proportion of redfish which I use to estimate the true proportion. Thus, here the margin of error M is given by

$$M = \frac{z\sigma_p}{\sqrt{n}},\ z = invNorm(.975, 0, 1),\ \sigma_p = \sqrt{p(1-p)} = \sigma_{I_A}.$$

Since $0 \leq p \leq 1$, and $\sigma_p = \sqrt{p(1-p)}$, it follows that

$$\sigma_p \leq \frac{1}{2}.$$

We therefore work this problem just like the previous problem using $B = 1/2$ and the .25 replaced by .02. The result is going to be roughly (since z is slightly less than 2),

$$n \geq \left(\frac{2 \cdot (1/2)}{.02}\right)^2 = \left(\frac{1}{.02}\right)^2 = (50)^2 = 2500,$$

that is in general, for a 95 percent confidence interval for a proportion, a conservative rough estimate that is easy, is simply take the inverse of the allowed margin of error and square it. To get the accurate answer, we calculate

$$n \geq \left(\frac{z \cdot (1/2)}{.02}\right)^2,\ z = invNorm(.975, 0, 1),$$

which gives, since $1/(.02) = 50$,

$$n \geq ((50)(.5) \cdot invNorm(.975, 0, 1))^2 = 2400.911766,$$

and therfore as n must be a whole number, we conclude

$$n \geq 2401.$$

FINAL ANSWER: 2401

86. Suppose that K and M are statements and that $P(K\&M) = .4$, and $P(M) = .7$. What is $P(K|M)$?

ANSWER: We know that

$$P(K|M) = \frac{P(K\&M)}{P(M)}.$$

Therefore

$$P(K|M) = \frac{.4}{.7} = \frac{4}{7} \text{ or about } .571.$$

FINAL ANSWER: 4/7 or .571

87. Suppose that K and M are statements and that $P(K\&M) = .4,\ P(K) = .8$, and $P(M) = .7$. What is $P(M|K)$?

ANSWER: We know that

$$P(M|K) = \frac{P(M\&K)}{P(K)},$$

so

$$P(M|K) = \frac{.4}{.8} = \frac{1}{2} \text{ or } .5.$$

FINAL ANSWER: 1/2 or .5

88. Suppose that K and M are statements and that $P(K|M) = 4/7,\ P(K) = .8$, and $P(M) = .7$. What is $P(M|K)$?

ANSWER: We know that

$$P(M|K) = \frac{P(M\&K)}{P(K)},$$

but $P(M\&K)$ is not directly given to us. However, we don know that

$$P(M\&K) = P(K\&M) = P(K|M)P(M),$$

and we are given $P(K|M)$ and $P(M)$. Therefore,

$$P(M\&K) = \frac{4}{7}(.7) = (\frac{4}{7})(\frac{7}{10}) = \frac{4}{10} = \frac{2}{5} \text{ or } .4$$

Now, we know that

$$P(M|K) = PM\&K)/P(K) = \frac{.4}{.8} = \frac{4}{8} = \frac{1}{2} = .5$$

FINAL ANSWER: 1/2 =.5

89. Suppose that K and M are statements and that $P(K|M) = .3,\ P(K|\text{not } M) = .8$, and $P(M) = .4$. What is $P(K)$?

ANSWER: We know that

$$P(K) = P(K\&M) + P(K\& \text{ not } M) = P(K|M)P(M) + P(K|\text{not } M)P(\text{not } M).$$

Since $P(\text{not } M) = 1 - P(M)$, this means

$$P(K) = P(K|M)P(M) + P(K|\text{not } M)[1 - P(M)].$$

Therefore

$$P(K) = (.3)(.4) + (.8)[1 - .4] = .12 + (.8)(.6) = .12 + .48 = .6 \text{ or } \frac{3}{5}.$$

FINAL ANSWER: 3/5 or .6

90. Suppose that K and M are statements and that $P(K|M) = .3,\ P(K|\text{not } M) = .8$, and $P(M) = .4$. What is $P(M|K)$?

ANSWER: We know that

$$P(M|K) = \frac{P(M\&K)}{P(K)}.$$

We also know that

$$P(M\&K) = P(K\&M) = P(K|M)P(M).$$

Therefore,

$$P(M\&K) = (.3)(.4) = .12.$$

On the other hand, from the previous problem, we know that $P(K) = .6$. Therefore,

$$P(M|K) = \frac{.12}{.6} = \frac{12}{60} = \frac{1}{5} \text{ or } .2.$$

You should notice here, that if you look back at the calculation in the previous problem, we also there calculated $P(M\&K) = .12$ in the process of calculating the first term that went into $P(K)$, so we could really just look back at the previous calculation and get the answer to this almost instantaneously.

FINAL ANSWER: 1/5 or .2

91. Suppose that K, L, M, N are statements, that exactly one of the three statements L, M, N is true, and that

$$P(K|L) = .7,\ P(K|M) = .4,\ P(K|N) = .6,\ P(L) = .3,\ P(M) = .5,\ P(N) = .2.$$

Find $P(K)$.

ANSWER: Here we have

$$P(K) = P(K\&L) + P(K\&M) + P(K\&N),$$

so

$$P(K) = P(K|L)P(L) + PK|M)P(M) + P(K|N)P(N).$$

Therefore

$$P(K) = (.7)(.3) + (.4)(.5) + (.6)(.2) = .21 + .2 + .12 = .53.$$

Notice that even though it was not asked for, in this process of calculating $P(K)$, we have also calculated (just look at the three terms we added):

$$P(K\&L) = .21,$$

$$P(K\&M) = .2,$$

$$P(K\&N) = .12,$$

which will be useful in the following problems.

FINAL ANSWER: .53

92. Suppose that K, L, M, N are statements, that exactly one of the three statements L, M, N is true, and that

$$P(K|L) = .7,\ P(K|M) = .4,\ P(K|N) = .6,\ P(L) = .3,\ P(M) = .5,\ P(N) = .2.$$

Find $P(L|K)$.

ANSWER: We know that

$$P(L|K) = \frac{P(L\&K)}{P(K)},$$

and from the previous problem, we know that

$$P(L\&K) = P(K\&L) = .21,$$

and

$$P(K) = .53,$$

so

$$P(L|K) = \frac{.21}{.53} = \frac{21}{53} \text{ or about } .396.$$

FINAL ANSWER: 21/53 or about .396

93. Suppose that K, L, M, N are statements, that exactly one of the three statements L, M, N is true, and that

$$P(K|L) = .7,\ P(K|M) = .4,\ P(K|N) = .6,\ P(L) = .3,\ P(M) = .5,\ P(N) = .2.$$

Find $P(M|K)$.

ANSWER: We know that

$$P(M|K) = \frac{P(M\&K)}{P(K)},$$

and from the previous problem, we know that

$$P(M\&K) = P(K\&M) = .2,$$

and

$$P(K) = .53,$$

so

$$P(L|K) = \frac{.2}{.53} = \frac{20}{53} \text{ or about } .377.$$

FINAL ANSWER: 20/53 or about .377

94. Suppose that K, L, M, N are statements, that exactly one of the three statements L, M, N is true, and that

$$P(K|L) = .7,\ P(K|M) = .4,\ P(K|N) = .6,\ P(L) = .3,\ P(M) = .5,\ P(N) = .2.$$

Find $P(N|K)$.

ANSWER: We know that

$$P(N|K) = \frac{P(N\&K)}{P(K)},$$

and from the previous problem, we know that

$$P(N\&K) = P(K\&N) = .21,$$

and

$$P(K) = .53,$$

so

$$P(L|K) = \frac{.12}{.53} = \frac{12}{53} \text{ or about } .226.$$

FINAL ANSWER: 12/53 or about .226

95. Suppose that Sam is taking a multiple choice test to get his driving license renewed. Each question has 5 possible answers to choose from and exactly one answer is correct. Sam knows the correct answers to 70 percent of the questions, but he is not feeling well, and he only marks the questions he knows correctly 90 percent of the time. Of course, if he does not know, he just guesses by marking a random selection.

What is the probability that Sam marks a question correctly?

ANSWER: It helps to define some symbols to represent the various statements here. Suppose that Q is a question on the test, but you do not know whether it is a question for which Sam knows the answer. Let K be the statement the Sam knows the answer to Q, and let C be the statement that Sam marks the correct answer to Q.. The information in the problem then tells us that

we need to find $P(C)$,

and it also tells us:

$P(C|K) = .9$
$P(K) = .7$
$P(C|\text{not } K) = 1/5 = .2.$

Now, we are in the situation of the previous problems. We know that

$$P(C) = P(C\&K) + P(C\& \text{ not } K)$$

and therefore

$$P(C) = P(C|K)P(K) + P(C| \text{ not } K)P(\text{ not } K).$$

Also, $P(\text{not } K) = 1 - P(K) = 1 - .7 = .3$.

Therefore,

$$P(C) = (.9)(.7) + (.2)(.3) = .63 + .06 = .69.$$

Notice again, in the process of calculating $P(C)$, we have also calculated

$$P(C\&K) = .63,$$

and

$$P(C\& \text{ not } K) = .06.$$

FINAL ANSWER: .69

96. Suppose that Sam is taking a multiple choice test to get his driving license renewed. Each question has 5 possible answers to choose from and exactly one answer is correct. Sam knows the correct answers to 70 percent of the questions, but he is not feeling well, and he only marks the questions he knows correctly 90 percent of the time. Of course, if he does not know, he just guesses by marking a random selection.

(a) What is the probability that Sam knows the answer to a question he marks correctly?

(b) What is the probability that Sam does not know the answer to a question he marks incorrectly?

ANSWER: From the previous problem, we have

$$P(C\&K) = .63,$$

$$P(C\& \text{ not } K) = .06,$$

and

$$P(C) = .69.$$

For the answer to (a) we must realize that it is asking for $P(K|C)$. We then know that

$$P(K|C) = \frac{P(K\&C)}{P(C)} = \frac{.63}{.69} = \frac{21}{23} \text{ or about } .913.$$

For the answer to (b), we must realize that it asks for $P(\text{not } K| \text{ not } C)$ which we know is

$$P(\text{ not } K| \text{ not } C) = 1 - P(K| \text{ not } C),$$

and

$$P(K| \text{ not } C) = \frac{P(K\&\text{not } C)}{P(\text{ not } C)} = \frac{P(C) - P(C\&K)}{1 - P(C)} = \frac{.69 - .63}{1 - .69} = \frac{6}{31},$$

so

$$P(\text{ not } K| \text{ not } C) = 1 - P(K| \text{ not } C) = 1 - \frac{6}{31} = \frac{25}{31} \text{ or about } .806.$$

FINAL ANSWER: (a) 21/23 or about .913 (b) 25/31 or about .806

97. Ball Point Stylus Pen Limited, LLC (BPSPLLLC), has three factories, one is in Beijing, one is in Singapore, and one is in Rangoon. The Beijing factory produces 60 percent of the ball point pens produced by BPSPLLLC, the Singapore factor produces 25 percent of the ballpoint pens produced by BPSPLLLC, and the Rangoon factory produces the remaining 15 percent of the ball point pens produced by BPSPLLLC. The CEO of BPSPLLLC is concerned about defective ballpoint pens in the output. He knows that at the Beijing factory, 5 percent of the output is defective, at the Singapore factory 8 percent of the output is defective, and at the Rangoon factory, 10 percent of the output is defective.

What is the overall percentage of defective ballpoint pens produced by BPSPLLLC, that is, what is the probability that a ballpoint pen is defective, given the only thing we know is that it was produced by BPSPLLLC.

ANSWER: Begin by introducing symbols to aid calculation. We suppose that we deal with a ballpoint pen which we only know is manufactured by BPSPLLLC. Let D be the statement that it is defective. Let B be the statement that it was made in Beijing, let G be the statement that it was made in Singapore, and let R be the statement that it was made in Rangoon. In terms of these symbols, if we can find $P(D)$, then expressed as a percentage, we know that it is the overall percentage of defective ballpoint pens produced by BPSPLLLC. In other words, we are really being asked for $P(D)$, but we must express our answer as a percentage.

The problem information tells us

$P(B) = .6,\ P(D|B) = .05,$

$P(G) = .25,\ P(D|G) = .08,$

$P(R) = .15$, and $P(D|R) = .1$.

Now, we know, using the same methods as in previous problems,

$$P(D) = P(D\&B) + P(D\&G) + P(D\&R),$$

so

$$P(D) = P(D|B)P(B) + P(D|G)P(G) + P(D|R)P(R).$$

Thus, using the formula $P(M\&N) = P(M|N)P(N) <$ for any statements M and N, we have

$P(D\&B) = .03,$

$P(D\&G) = .02,$

$P(D\&R) = .015.$

Thus,

$$P(D) = .03 + .02 + .015 = .065,$$

Which means that the overall percentage of defective ballpoint pens produced by BPSPLLLC is 6.5 percent.

FINAL ANSWER: 6.5 percent

98. Ball Point Stylus Pen Limited, LLC (BPSPLLLC), has three factories, one is in Beijing, one is in Singapore, and one is in Rangoon. The Beijing factory produces 60 percent of the ball point pens produced by BPSPLLLC, the Singapore factor produces 25 percent of the ballpoint pens produced by BPSPLLLC, and the Rangoon factory produces the remaining 15 percent of the ball point pens produced by BPSPLLLC. The CEO of BPSPLLLC is concerned about defective ballpoint pens in the output. He knows that at the Beijing factory, 5 percent of the output is defective, at the Singapore factory 8 percent of the output is defective, and at the Rangoon factory, 10 percent of the output is defective.

Given that a BPSPLLLC ballpoint pen is defective, what is the chance it was produced in the Beijing factory?

ANSWER: Begin by introducing symbols to aid calculation. We suppose that we deal with a ballpoint pen which we only know is manufactured by BPSPLLLC. Let D be the statement that it is defective. Let B be the statement that it was made in Beijing, let G be the statement that it was made in Singapore, and let R be the statement that it was made in Rangoon. Here we are being asked for $P(B|D)$.

The problem information tells us

$P(B) = .6,\ P(D|B) = .05,$

$P(G) = .25,\ P(D|G) = .08,$

$P(R) = .15$, and $P(D|R) = .1$.

Now, we know, using the same methods as in previous problems how to proceed and from the caculations in the previous problem, that is using

$$P(M\&N) = P(M|N)P(N), \text{ for any statements } M,\ N,$$

we have

$P(D\&B) = .03,$

$P(D\&G) = .02,$

$P(D\&R) = .015.$

We know that

$$P(B|D) = \frac{P(B\&D)}{P(D)} = \frac{.03}{.065} = \frac{30}{65} = \frac{6}{13} \text{ or about } .462.$$

FINAL ANSWER: 6/13 or about .462

99. Ball Point Stylus Pen Limited, LLC (BPSPLLLC), has three factories, one is in Beijing, one is in Singapore, and one is in Rangoon. The Beijing factory produces 60 percent of the ball point pens produced by BPSPLLLC, the Singapore factor produces 25 percent of the ballpoint pens produced by BPSPLLLC, and the Rangoon factory produces the remaining 15 percent of the ball point pens produced by BPSPLLLC. The CEO of BPSPLLLC is concerned about defective ballpoint pens in the output. He knows that at the Beijing factory, 5 percent of the output is defective, at the Singapore factory 8 percent of the output is defective, and at the Rangoon factory, 10 percent of the output is defective.

Given that a BPSPLLLC ballpoint pen is defective, what is the chance it was produced in the Singapore factory?

ANSWER: Begin by introducing symbols to aid calculation. We suppose that we deal with a ballpoint pen which we only know is manufactured by BPSPLLLC. Let D be the statement that it is defective. Let B be the statement that it was made in Beijing, let G be the statement that it was made in Singapore, and let R be the statement that it was made in Rangoon. Here we are being asked for $P(G|D)$.

The problem information tells us

$P(B) = .6,\ P(D|B) = .05,$

$P(G) = .25,\ P(D|G) = .08,$

$P(R) = .15$, and $P(D|R) = .1$.

Now, we know, using the same methods as in previous problems how to proceed and from the caculations in the previous problem, that is using

$$P(M\&N) = P(M|N)P(N),\ \text{for any statements } M,\ N,$$

we have

$P(D\&B) = .03,$

$P(D\&G) = .02,$

$P(D\&R) = .015.$

We know that

$$P(G|D) = \frac{P(G\&D)}{P(D)} = \frac{.02}{.065} = \frac{20}{65} = \frac{4}{13} \text{ or about } .308.$$

FINAL ANSWER: 4/13 or about .308

100. Ball Point Stylus Pen Limited, LLC (BPSPLLLC), has three factories, one is in Beijing, one is in Singapore, and one is in Rangoon. The Beijing factory produces 60 percent of the ball point pens produced by BPSPLLLC, the Singapore factor produces 25 percent of the ballpoint pens produced by BPSPLLLC, and the Rangoon factory produces the remaining 15 percent of the ball point pens produced by BPSPLLLC. The CEO of BPSPLLLC is concerned about defective ballpoint pens in the output. He knows that at the Beijing factory, 5 percent of the output is defective, at the Singapore factory 8 percent of the output is defective, and at the Rangoon factory, 10 percent of the output is defective.

Given that a BPSPLLLC ballpoint pen is defective, what is the chance it was produced in the Rangoon factory?

ANSWER: Begin by introducing symbols to aid calculation. We suppose that we deal with a ballpoint pen which we only know is manufactured by BPSPLLLC. Let D be the statement that it is defective. Let B be the statement that it was made in Beijing, let G be the statement that it was made in Singapore, and let R be the statement that it was made in Rangoon. Here we are being asked for $P(R|D)$.

The problem information tells us

$P(B) = .6,\ P(D|B) = .05,$

$P(G) = .25,\ P(D|G) = .08,$

$P(R) = .15$, and $P(D|R) = .1$.

Now, we know, using the same methods as in previous problems how to proceed and from the caculations in the previous problem, that is using

$$P(M\&N) = P(M|N)P(N),\ \text{for any statements } M,\ N,$$

we have

$P(D\&B) = .03,$

$P(D\&G) = .02,$

$P(D\&R) = .015.$

We know that

$$P(R|D) = \frac{P(R\&D)}{P(D)} = \frac{.015}{.065} = \frac{15}{65} = \frac{3}{13} \text{ or about } .231.$$

FINAL ANSWER: 3/13 or about .231