Statistics Practice

AP Statistics Exam Review #1

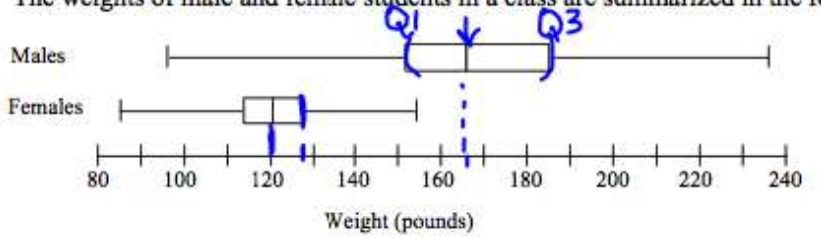
C Q

1. Determine whether the given variable is categorical or quantitative. If it is quantitative, indicate if it is continuous or discrete.
   a. Score on a quiz for a class (out of 100 points). Q, discrete
   b. Final grade for a course (A, B, C, D, F). C
   c. The time it takes to be in line at the driver's license office. Q, continuous (interval)
   d. The number of classes a student missed. Q, discrete.
   e. The heights of all professional basketball players. Q, continuous
   f. The different makes of cars in the school parking lot. C
2. Of the range, the interquartile range, and the variance, which is most influenced by an outlying value in the data set? → influenced
   ↳ Q3 − Q1 ⇒ IQR not influenced
   Range
3. A data set has only positive values. If the largest value of a data set is doubled, which of the following is not true?
   a. The mean increases. ✓
   b. The range increases. ✓
   c. The interquartile range increases. ✗
   d. The standard deviation increases. ✓
   e. All of these are true.

   2, 3, 7, 10
   new: 2, 3, 7, 20
4. If the test scores of a class of 30 students have a mean of 75.6 and the test scores of another class of 24 students have a mean of 68.4, what is the mean of the combined group?

   $[30(75.6) + 24(68.4)]/54$
   $= 72.4$
5. If a distribution has zero variance, which of the following is true?
   a. All the values are positive.
   b. All the values are negative.
   c. The number of positive values and the number of negative values are equal.
   d. All the values are equal to each other.
6. A set of data is found to have a sample standard deviation of 25. Suppose that 6 is added to each of the numbers in the data set.
   a. What can you say about the standard deviation of the new data set? 25
   b. What can you say about the mean of the new data set? increased by 6

7. The weights of male and female students in a class are summarized in the following boxplots:



Which of the following is NOT correct?

a. The male students have less variability than the female students. False

b. About 50% of the male students have weights between 150 and 185 lbs. True

c. About 25% of the female students have weights more than 128 lbs. True

d. The median weight of the male students is about 166 lbs. True

e. The mean weight of the female students is about 120 because of symmetry. True
(median)

8. The following is a stem-plot of the birth weights of male babies born to the smoking group. The stems are in units of kg.

n = 26

Stems	Leaves
2	3,4,6,7,7,8,8,8,9
3	2,2,3,4,6,7,8,9
4	1,2,2,3,4
5	3,5,5,6

median = $\frac{3.4+3.6}{2} = 3.5$

a. What is the median birth weight of the data?

b. Change this into a split stem-plot.

c. What can be said about the shape of the graphs?

Skewed right

Stem	
2	3, 4
2	6, 7, 7, 8, 8, 8, 9
3	2, 2, 3, 4
3	6, 7, 8, 9
4	1, 2, 2, 3, 4
4	
5	3
5	5, 5, 6

9. A survey was conducted to gather ratings of the quality of service at local restaurants. Respondents rated on a scale of 0 (terrible) to 100 (excellent). The data are represented by the following stem plot.

3	2 4
4	0 3 4 7 8 9 9 9
5	0 1 1 2 3 4 5
6	1 2 5 6 6
7	0 1
8	
9	2

stem = tens
leaves = ones

a. Find the mean and standard deviation of this data. $\bar{x} = 54.16 \quad s = 12.928$

b. Create a box plot from the data.

minX=32
Q1=47.5
Med=51
Q3=63.5
maxX=92

IQR = 63.5 − 47.5 = 16.0

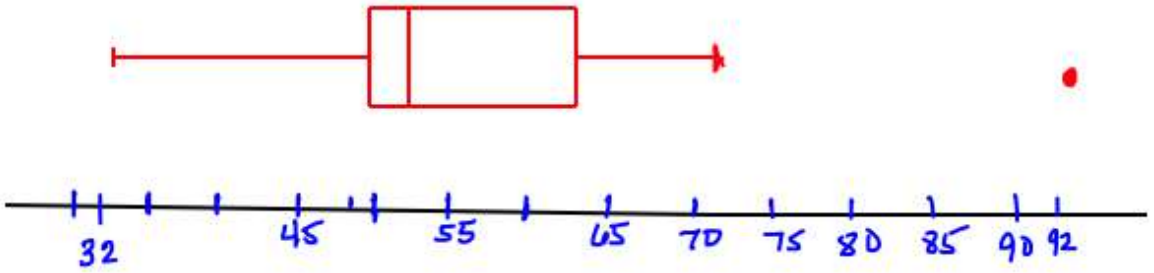
1.5(IQR) = 24

Outlier boundaries: Q1 − 1.5IQR = 47.5 − 24 = 23.5

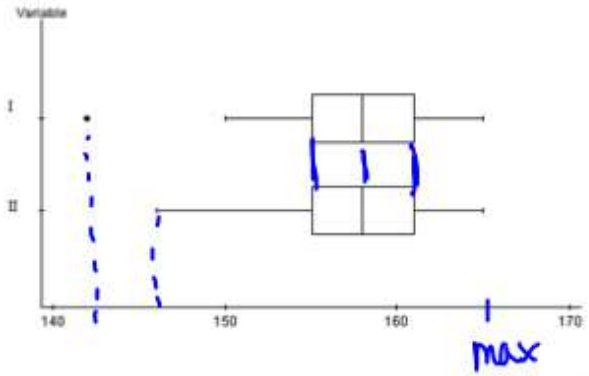
Q3 + 1.5IQR = 63.5 + 24 = 87.5

[23.5, 87.5]

92 is an outlier



10. The boxplots shown below summarize two data sets, I and II. Based on the boxplots, which of the following statements about these two data sets CANNOT be justified?



a. Data set I and data set II have the same number of data points. False

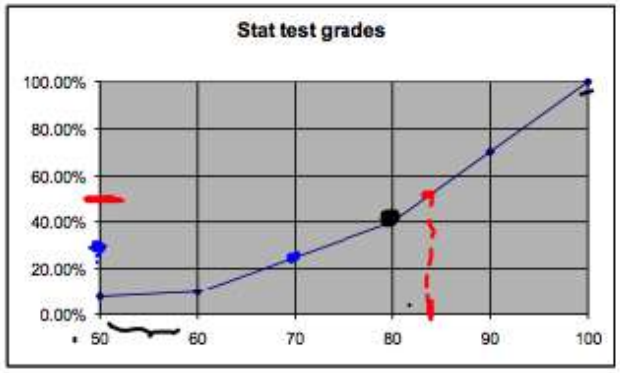
b. The range of data set I is greater than the range of data set II.

c. The interquartile range of data set I is equal to the interquartile range of data set II.

d. The median of data set I is equal to the median of data set II.

e. All of the above are valid statements.

11. The figure below shows a **cumulative** relative frequency histogram of 40 scores on a test given in an AP Statistics class. Which of the following conclusions can be made from the graph?



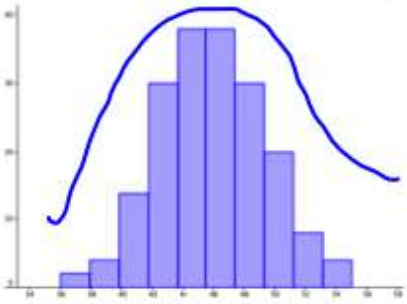
F a. If the passing score is 70, most students did not pass the test. 70 ≈ 25th percentile

F b. The median test score is less than 70.

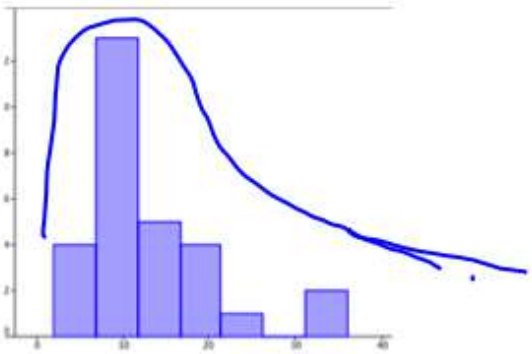
T c. Sixty percent of the students had a test score above 80.

F d. The horizontal nature of the graph for test scores of 60 and below indicates that those scores occurred most frequently.

12. Determine if the following graphs are skewed right, skewed left or roughly symmetrical:

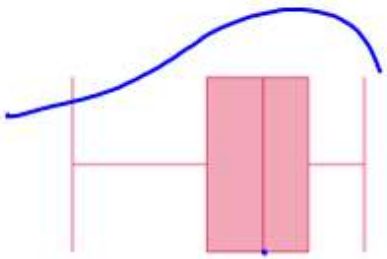
a. 

Symmetrical

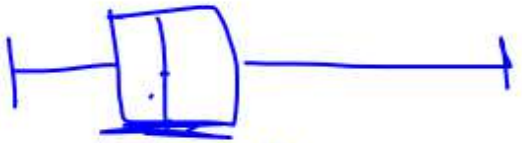
b. 

skewed right

skewed left

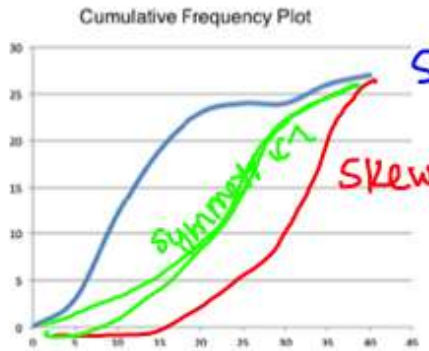


skewed left



skewed right

c. 



skewed right

skewed left

FOR BINOMIAL PROBABILITIES (Pg 387):

You need to know how to find the number of combinations

See green sheet: $P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$ where X = R.V. n = number trials, k = the k^{th} success

19. In how many ways can a committee of 3 be chosen from a group of 7 people? $nCr = 7C3$ "7 CHOOSE 3"

$= 35$

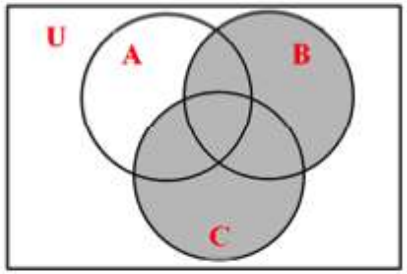
CALC: $\binom{n}{k} = \binom{7}{3} =$ 7 (1) nCr (2) (MATH PRB 3:nCr) 3 (3) $= 35$

Permutation - order matters

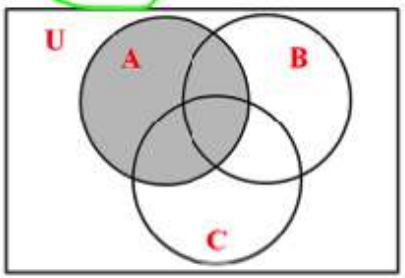
↳ NOT ON AP EXAM

20. Which of the following is the correct Venn diagram for $A \cap (B \cup C)$?

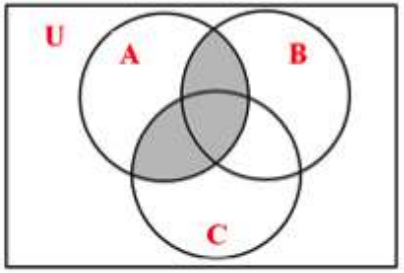
a.



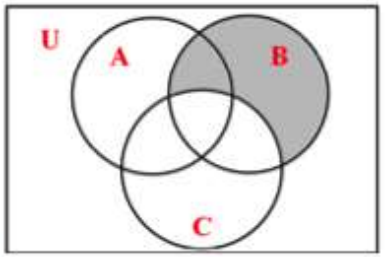
c.

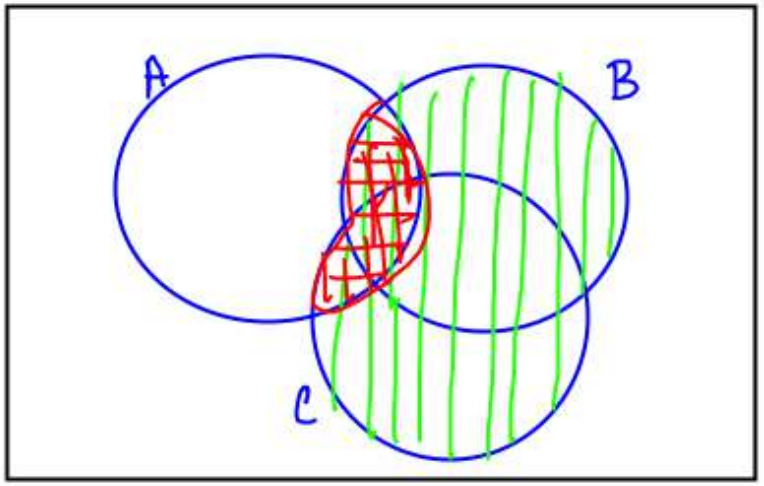


b.



d.





21. Given:

$U = \{1,2,3,4,5,6,7,8,9,10\}$

$A = \{2,4,6,8,10\}$

$B = \{1,2,3,4\}$

$C = \{4,5,6,9\}$

a. Find B^C
b. Find $A \cap B$
c. Find $A \cap B \cap C$
d. Find $A \cap (B \cup C)^C$

empty set
{} or $\emptyset$

a. $B^C = \{5,6,7,8,9,10\}$

b. $A \cap B = \{2,4\}$

c. $A \cap B \cap C = \{4\}$

d. $A \cap (B \cup C)^C$

$B \cup C = \{1,2,3,4,5,6,9\}$

$(B \cup C)^C = \{7,8,10\}$

$A \cap (B \cup C)^C = \{8,10\}$

22. If the probabilities that Jane, Tom, and Mary will be chosen chairperson of the board are 0.5, 0.3 and 0.2 respectively, what is the probability that the chairperson will be either Jane or Mary?

Mutually exclusive
⇒ intersection is ∅

$.5 + .2 = .7$

23. Suppose a box contains 3 defective light bulbs and 12 good bulbs. Two bulbs are chosen from the box without replacement. Find the probability that one of the bulbs drawn is good and one is defective.

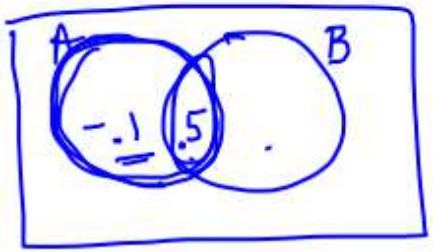
$n = 15$ bulbs

$n(S) = {}_{15}C_2 = 105$

$P(\text{one good and one defective}) = \frac{{}_3C_1 \cdot {}_{12}C_1}{105} = \frac{36}{105} = \frac{12}{35}$

24. Identify why this assignment of probabilities cannot be legitimate:
→ P(A) = .4, P(B) = .3, P(A and B) = .5

a. A and B are not given as mutually exclusive events
b. P(A and B) cannot be greater than either P(A) or P(B).
c. P(A|B) is not known.
d. A and B are not given as independent events.
e. The assignment is legitimate.
f. None of these



25. Which of the following pairs of events are mutually exclusive (disjoint)?

a. Choosing a King and a Heart from a standard deck of cards
b. Studying French and Spanish
c. Choosing a King and an Ace from a standard deck of cards
d. The events of rolling a sum of 6 and of rolling a double on a pair of die (number cubes)
e. No pair is mutually exclusive
f. All of these examples are mutually exclusive

33

Independent $\Rightarrow P(A)\cdot P(B) = P(A \cap B)$ $\quad P(A|B) = \frac{P(A \cap B)}{P(B)}$

$P(A|B) = P(A)$

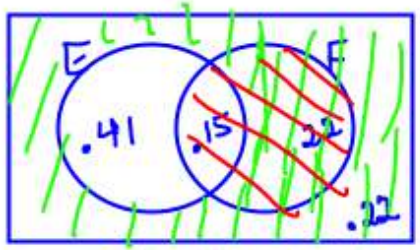
26. Given:

$$P(E)=0.56, P(F)=0.37, P(E \cup F)=0.78$$

$P(E \cup F) = P(E) + P(F) - P(E \cap F)$

$.78 = .56 + .37 - P(E \cap F)$

a. $P(E \cap F) =$ .15
b. $P(E^C \cup F) =$ .59
c. $P(E|F) =$
d. $P(F|E) =$
e. Are E and F independent? NO



c. $P(E|F) = .15/.37 = 15/37 = .4054\ldots$

d. $P(F|E) = .15/.56 = 15/56$

27. If sets A and B are independent and $P(A) = .36$ and $P(B) = .58$, find $P(A \cup B)$.

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$= .36 + .58 - [(.36)(.58)] = .7312$

D = defective

28. A VCR manufacturer receives 60% of his parts from factory F1 and the rest from factory F2. Suppose that 3% of the outputs from F1 are defective while only 2% of the outputs from F2 are defective.

    a. What is the probability that a received part is defective?

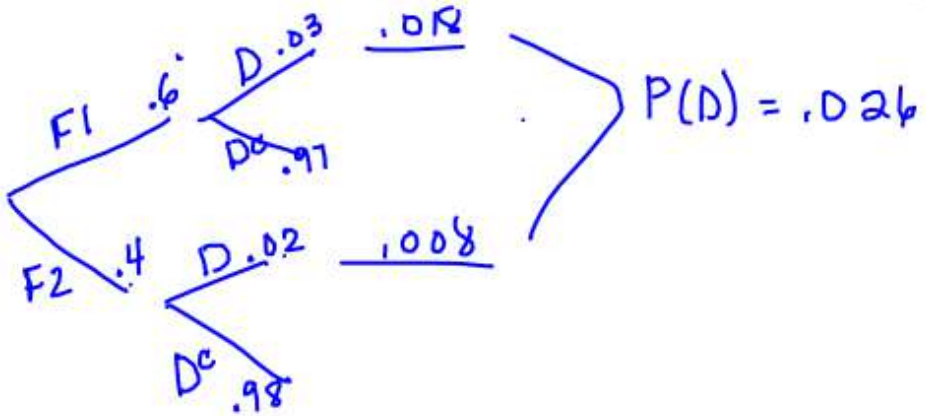
    b. If a randomly chosen part is defective, what is the probability it came from factory F1?

9/13

$P(F1|D) = \frac{P(F1 \cap D)}{P(D)} = \frac{.018}{.026}$

$P(D|F1) = .03$

$P(F1) = .6$



$P(D) = .026$

$P(D|F1) = \frac{P(D \cap F1)}{P(F1)}$

$.03 = \frac{P(D \cap F1)}{.6}$

$P(D \cap F1) = .018$

$P(D \cap F2) = .008$

$\boxed{P(D) = .026}$

29. Two different airlines have a flight from LA to New York that departs each weekday morning at a certain time. Suppose that the probability that the first airline is fully booked on a particular day is .7, the probability that the second airline is fully booked on a particular day is .6 and the probability that they are both booked is .54.

    a. Calculate the probability that exactly one of the airlines is booked. $.16 + .06 = .22$

    b. Calculate the probability that at least one of the airlines is booked. .76

    c. Calculate the probability that the second airline is booked, given that the first airline is booked. $P(B|A) = \frac{P(A \cap B)}{P(A)} = .54/.7 = 27/35$

    d. Calculate the probability that the first airline is booked, given that the second airline is booked. $P(A|B) = \frac{P(A \cap B)}{P(B)} = .54/.6 = 9/10$

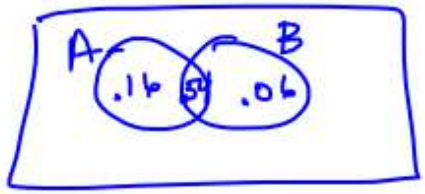
A = 1st airline booked

B = 2nd airline booked

$P(A) = .7$

$P(B) = .6$

$P(A \cap B) = .54$



30. Using a standard deck of 52 cards,

a. how many different 5 card hands are possible? $_{52}C_5 = 2598960$

b. what is the probability of getting a full house (three of a kind and two of a kind within a five card hand)?

c. what is the probability of getting four of a kind in a five card hand?

Example set of 52 poker playing cards

Suit	Ace	2	3	4	5	6	7	8	9	10	Jack	Queen	King
Clubs													
Diamonds													
Hearts													
Spades													

b.

$13\,(_4C_3)\cdot 12\,(_4C_2) = 3744$

$P(\text{full house}) = \frac{3744}{2598960} \approx .00144$

c.

$13\,(_4C_4)\cdot 12\,(_4C_1) = 624$

$P(\text{4 of a kind}) = \frac{624}{2598960} \approx .00024$

Statistics Practice

AP Statistics Exam Review #2

I. Bivariate Data

1. A value of the correlation coefficient equal to ___-1___ indicates that there is a perfect negative relationship between the dependent variable y and the independent variable x.

2. A positive correlation coefficient indicates that large values of x are associated with ___large___ values of y.

3. In regression analysis, the variable that is being predicted is called the ___response___ variable.

4. Generally, the larger the correlation (either positive or negative) between the variables for the simple linear regression model, the (better, worse) ___better___ will be the predictions of y for given values of x.

5. For the simple linear regression model, if all the points on a scatter plot lie in a straight line with correlation coefficient r = -1, then the slope of the regression line is
   a. -1
   b. +1
   c. positive
   d. negative

6. True or False: The coefficient of determination can assume negative values.

Use the following set of observations for the independent variable x and the dependent variable y in questions #7-9:

X	-3	-1	1	3
Y	8	4	5	-1

7. The correlation coefficient is : $r = -.8971$

8. The coefficient of determination is: $r^2 = .8048$

9. The least squares linear regression equation is $y = 4 - 1.3x$

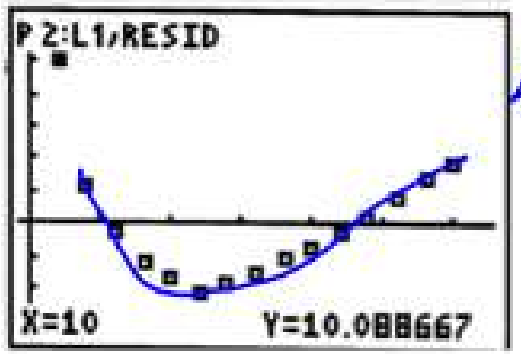
10. The correlation between two scores X and Y equals 0.8. If both the X scores and the Y scores are converted to z-scores, then the correlation between z-scores for X and z-scores for Y would be $r = .8$

11. In the least-squares regression line, the desired sum of the errors (residuals) should be 0

12. Suppose that you have a least squares regression line of $\hat{y} = 1.65 - 2.2x$ and that the actual y value corresponding to $x = 10$ is -19. What is the residual value corresponding to $y = -19$?
$res = y - \hat{y} = -19 - [1.65 - 2.2(10)] = 1.35$

13. The equation of the least squares regression line for a set of given points is $\hat{y} = 1.3 + 0.73x$. What is the residual for the point (4, 7)?
$res = 7 - \hat{y}(4) = 2.78$ x y

14. A study of fuel economy for various automobiles plotted the fuel consumption vs. speed. A LSRL was fit to the data. Here is the residual plot from this least-squares fit. What does the pattern of the residuals tell you about the linear model?

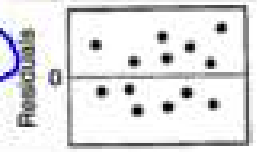


Contradicts the linear model

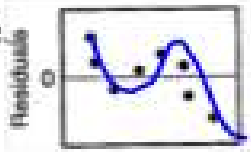
15. A prediction of the world's population in the year 2088 is an example of extrapolation.

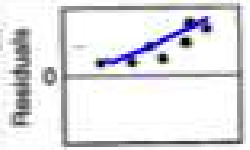
16. An observation that causes the values of the slope and the intercept in the line of best fit to be considerably different from what they would be if the observation were removed from the data set is said to be influential.

17. Which of the following residual plots indicates a reasonable fit to a given set of data?

a. 

b. 

c. 

d. 

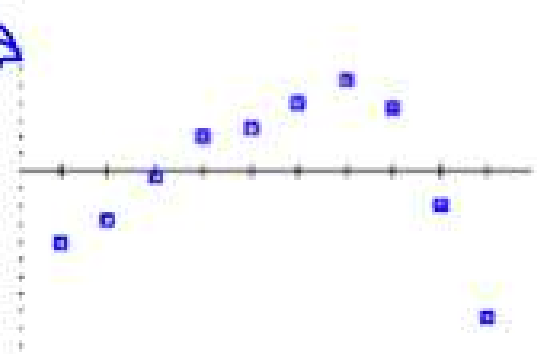
18. The following data lists the temperature readings (in degrees Fahrenheit) at different settings on a thermostat for an experimental cooling container.

x	Setting	1	2	3	4	5	6	7	8	9	10
y	Temperature	88	85	83	81	77	74	71	65	55	44

a. What is the least squares regression line equation? $\hat{y} = 96.60 - 4.418x$

b. What is the value of the correlation coefficient? $r = -.9517$

c. Find the residual plot.



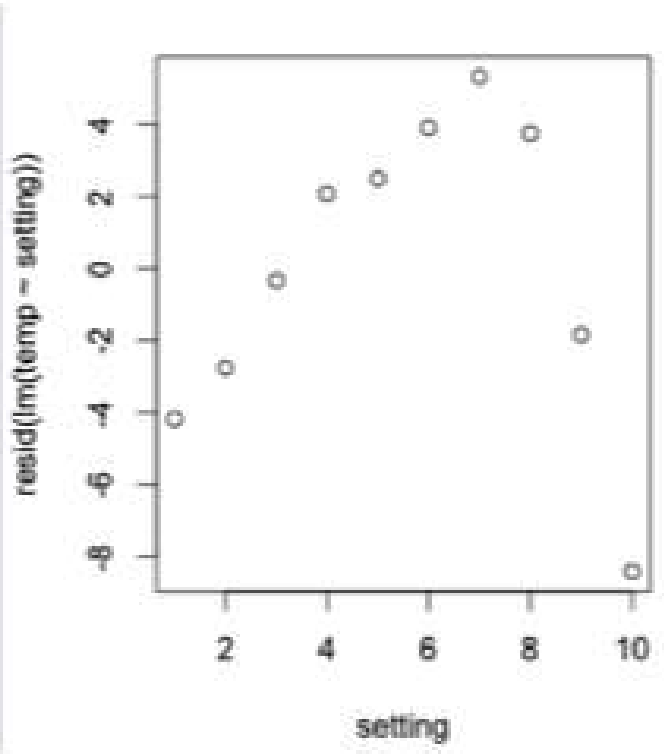
d. Would you say the LSRL is a good model? No

```
> setting
 [1]  1  2  3  4  5  6  7  8  9 10
> temp
 [1] 88 85 83 81 77 74 71 65 55 44
> lm(temp~setting)

Call:
lm(formula = temp ~ setting)

Coefficients:
(Intercept)      setting
     96.600       -4.418

> resid(lm(temp~setting))
         1          2          3          4          5
-4.1818182 -2.7636364 -0.3454545  2.0727273  2.4909091
         6          7          8          9         10
 3.9090909  5.3272727  3.7454545 -1.8363636 -8.4181818
> plot(setting,resid(lm(temp~setting)))
>
```



II. Sampling and Experiments

18. A marketing company offers to pay $25 to the first 100 persons who respond to their advertisement and complete a questionnaire regarding displays of their client's product. This situation is an example of what kind of sample? Voluntary Response

19. A SRS was selected of large urban school districts throughout New England. The selected districts were identified as target districts. Within each district, a SRS of its high schools was chosen and the principals of those high schools were interviewed. This situation is an example of a multistage sampling

20. In order to assess the membership's attitudes about a new Supreme Court decision, a local bar association selects a SRS of 100 lawyers from its membership list. Surveys are delivered to the selected lawyers. 63 of the lawyers return their surveys. 43 of the respondents disagree with the new ruling.
Which of the following is of great concern in this situation?
    a. Nothing is known about the parameters of the population of interest.
    b. Nothing is stated about the methodology of the SRS.
    c. There may be a problem with the sampling frame.
    d. There may be a problem with non-response bias.
    e. None of these is true.

True or False?

21. Voluntary response samples often under represent people with strong opinions. False

22. Convenience samples often lead to under coverage bias. True

23. Questionnaires with non-neutral wording are likely to have response bias. True

24. In an observational study we impose a treatment on the subjects. False

25. The entire group of individuals we want information about is called the sample. False

26. Which of the following describes a simple random sample (SRS)?
    a. successively smaller groups are selected within the population in stages
    b. choosing the individuals easiest to reach
    c. selecting one random value then choosing a cluster of subjects around it
    d. every possible sample of a given size has the same chance to be selected
    e. none of these

Statistics Practice

AP Statistics Exam Review #3

I. Discrete Random Variables

1. A distribution of grades in an introductory statistics class (where A = 4, B = 3, etc) is:

X	0	1	2	3	4
P(X)	.10	.25	.30	??	.10

a. Find $P(X = 3)$

b. Find $P(1 \le X < 3)$

c. Find the expected (mean) grade in this class.

d. Find the standard deviation for the class grades.

e. Find the lowest grade X_0 such that $P(X \ge X_0) < 0.5$

$E[X^2] = 0 + .32 + .32 + 1.62 + 4.8$
$= 7.06$

2. A discrete variable X has the following probability distribution:

X^2 0 1 4 9 16

X	0	1	2	3	4
P(X)	0.12	0.32	0.08	0.18	.3

a. Find $P(X = 4)$ $1 - (.12 + .32 + .08 + .18) = 1 - .7 = .3$

b. Find $P(1 < X \leq 3)$ $= .08 + .18 = .26$

$P(2 \leq X \leq 3)$

c. Find the mean grade in this class. $E[X] = \mu = 0 + .32 + .16 + .54 + 1.2 = 2.22$

d. Find the variance for the class grades. $Var[X] = E[X^2] - (E[X])^2 = 7.06 - (2.22)^2$
$= 2.1316$

e. Define a new variable Y such that Y = 3 X – 1. What are the mean and variance of Y?

$\mu_Y = E[Y] = E[3X - 1] = 3E[X] - 1 = 3(2.22) - 1 = 5.66$

$\sigma_Y^2 = Var[Y] = Var[3X - 1] = 9 \cdot Var[X] = 9(2.1316) = 19.1844$

$p = .8$ $n = 100$

3. A headache remedy is said to be 80% effective in curing headaches caused by simple nervous tension. An investigator tests this remedy on 100 randomly selected patients suffering from nervous tension.

   a. Define the random variable being measured. $X =$ the number of people who have headaches cured

   b. What kind of distribution does X have? binomial

   c. Calculate the mean of X $np = 80$

   d. Calculate the standard deviation of X $\sqrt{np(1-p)} = \sqrt{100(.8)(.2)} = 4$

   e. What is the probability that between 75 and 90 (inclusive) of the patients will obtain relief?

$P(75 \le X \le 90) = P(X \le 90) - P(X \le 74) = .9102$

0, 1, ... 73, 74, 75, ..., 89, 90, 91, ... 100

Geometric

$p = .6$ $1-p = .4$

4. A quarter back completes 60% of his passes. We want to observe this quarterback during one game to see how many pass attempts he makes before completing one pass.

   a. What is the probability that the quarterback throws 3 incomplete passes before he has a completion?

$$(.4)(.4)(.4)(.6) = .0384$$

   b. How many passes can the quarterback expect to throw before he completes a pass? (Round to nearest whole number)

$$E[X] = \frac{1}{.6} = \frac{10}{6} \approx 2$$

   c. Determine the probability that he makes more than 3 incomplete passes before he has a completion.

$$P(X>3) = (.4)^3 = .064$$

   d. What is the probability that he has 4 or fewer incomplete passes before he completes one?

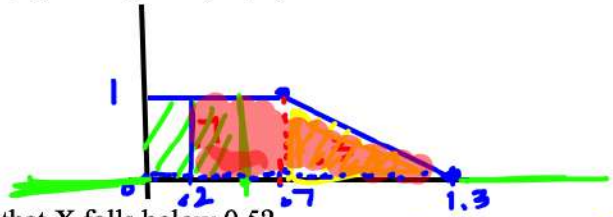
$$.6 + (.4)(.6) + (.4)^2(.6) + (.4)^3(.6) + (.4)^4(.6)$$

$$= .9897\overline{6}$$

geometcdf(.6, 5) or $1 - P(X>5) = 1 - .4^5$

II. Continuous Random Variables.

6. Think about a density curve that consists of two straight-line segments. The first goes from the point (0, 1) to the point (0.7, 1) then the second goes from the point (0.7, 1) to the point (1.3, 0)



.7 + .3 = 1 ✓

a. What is the probability that X falls below 0.5?

$P(X < .5) = (.5)(1) = .5$

b. What is the probability that X lies between 0.7 and 1.3?

$P(.7 < X\ 1.3) = .3$

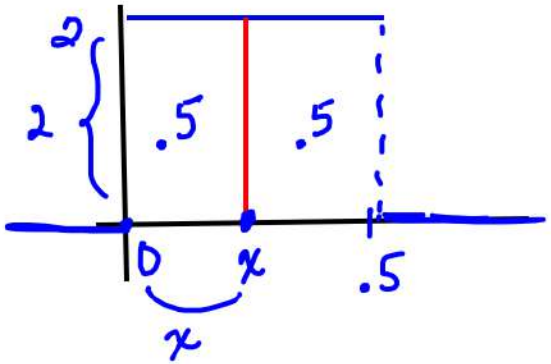
c. What is the probability that X lies above 0.2?

$P(X > .2) = (.5)(1) + .3 = .8$

d. What is the probability that X is equal to 0.5?

$P(X = .5) = 0$

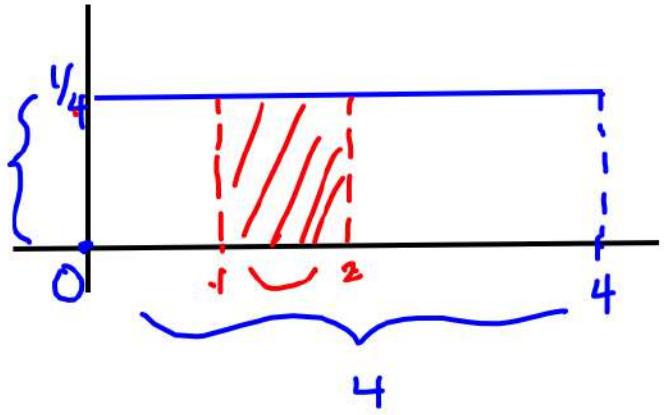
7. Suppose a density curve is defined by the function $f(x)= 2$ for all values between $x = 0$ to $x = 0.5$. Find the median of this density curve.



$.5 = 2x$

$.25 = x$

8. Consider a uniform distribution that is defined for $0 \le X \le 4$. What is the probability that X lies between 1 and 2? (Hint: draw a picture)



$4x = 1$

$x = 1/4$

$P(1 < X < 2) = 1/4 = .25$

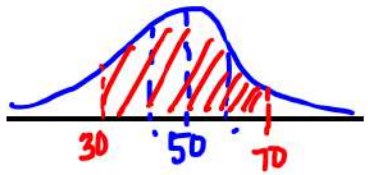
$N(50, 10)$

9. The distribution of raw scores on a chemistry final is approximately normal with mean of about 50 and standard deviation of about 10.

   a. What is the probability that a student will have a raw score that is more than 45?

$P(X > 45) = \text{normalcdf}(45, 9999999, 50, 10)$
$= .6915$

   b. Give the interval for the middle 95% of scores.



$[30, 70]$

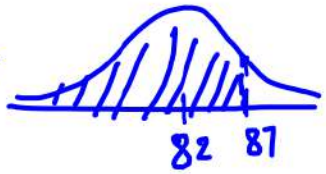
10. If a group of students have test scores that are normally distributed with a mean of 82 and a standard deviation of 4,

   a. half of the students made below a grade of __

82

   b. If a student made a grade of 87, what is their percentile rank? ~ 89th %

$P(X < 87) = .894$



   c. Find the probability that someone made less than a grade of 74.

$P(X < 74) = .02275$

$$z = \frac{x - \mu}{\sigma}$$

11. If a sample has a mean of 48 and a standard deviation of 3.2, what is the value in the set that corresponds to a z-score of -1.6?

(normal) μ σ x

$$-1.6 = \frac{x - 48}{3.2}$$

$x = 42.88$

12. The z-score associated with a measurement x represents the number of ____ that x lies above the mean or below the mean.

Standard deviations

13. Suppose family income in a particular suburb has been found to be approximately normal with a mean of \$52,137 and a standard deviation of \$19,452. What percentage has an income in the range of \$50,000 to \$80,000?

46.77%

$P(50000 < x < 80000) = \text{normalcdf}(50000, 80000, 52137, 19452)$

$= .4677$

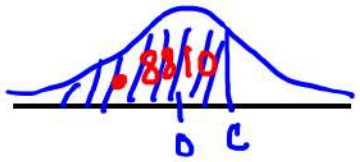
14. Find $P(-1.2 < Z < 1.9)$ = .8562

15. Find the value of c so that

a. $P(Z < c) = 0.8810$

invNorm(.8810)

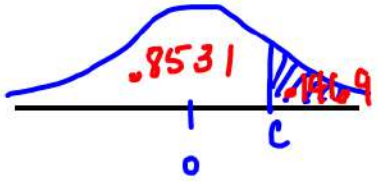
$c = 1.18$



b. $P(Z > c) = 0.1469$

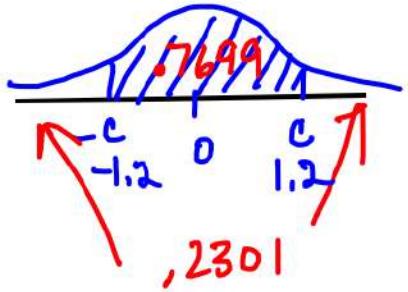
invNorm(.8531)

$c = 1.05$



c. $P(-c < Z < c) = 0.7699$

$c = 1.20$



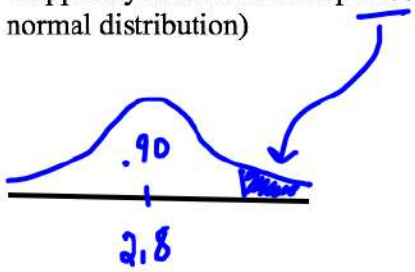
each tail is $\frac{.2301}{2} = .11505$

$P(Z < c) = .11505 + .7699$

$= .88495$

invNorm(.88495)

16. Suppose you rank in the top 10% of your class. If the mean GPA is 2.8 and the standard deviation is 0.43, what is your GPA? (Assume a normal distribution)



$\mu = 2.8$ $\sigma = .43$

$\text{InvNorm}(.90, 2.8, .43) = \boxed{3.35}$

$\text{InvNorm}(.9) = 1.282 \leftarrow$ z-score: $1.282 = \frac{x - 2.8}{.43}$

17. If a population has a standard deviation σ, then the standard deviation of the mean of 100 randomly selected items from this population is

$n = 100$

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{100}} = \frac{\sigma}{10}$

18. A waiter estimates that his average tip per table is \$20 with a standard deviation of \$4. If we take samples of 9 tables at a time, calculate the following probabilities when the tip per table is normally distributed.

$\mu_{\bar{x}} = 20$ $\sigma_{\bar{x}} = \frac{4}{\sqrt{9}}$ $n = 9$

a. What is the probability that his average tip is more than \$19?

$P(\bar{x} > 19) = \text{normalcdf}(19, 10000000, 20, 4/3)$
$= .7734$

b. What is the probability that his average tip is between \$19 and \$22?

$P(19 < \bar{x} < 22) = .7066$

$p = .86$

19. In a large population, 86% of the households own computers. A simple random sample of 100 households is to be contacted and the sample proportion computed.

$n = 100$

a. What is the mean and standard deviation of the sampling distribution of the sample proportions?

$\mu_{\hat{p}} = .86$ $\quad \sigma_{\hat{p}} = \sqrt{\frac{.86(1-.86)}{100}} = .0347$

b. What is the probability $\hat{p}$ will be between 82% and 90%?

$P(.82 < \hat{p} < .90) = \text{normalcdf}(.82, .90, .86, .0347)$

$= .751$

20. What is the difference between a discrete random variable and a continuous random variable?

Countable

interval

$P(X = \square) = 0$

Binomial

$p = .15$ $\quad n = 10$

5. 15% of the families in a certain subdivision have a pool in their back-yard. Suppose a sample of 10 families are taken at random from this subdivision.

   a. What is the probability that exactly 3 of the 10 families have a pool?

   $P(X = 3) = .1298$ $\quad$ binompdf(10, .15, 3)

   b. What is the probability that at least 2 of the families have a pool in their back-yard?

   $P(X \geq 2) = 1 - P(X \leq 1) = .04997$

   $= 1 - .5443 = .4557$

   (1 - binomcdf(10, .15, 1))

   c. What is the expected number of pools we should have from this sample?

   $E[X] = np = 10(.15) = 1.5$