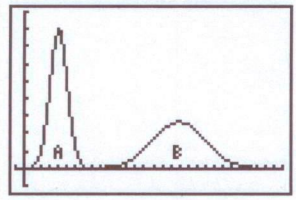
AP Statistics ONE LAST REVIEW PACKET! Name: key Per: __

1. **In the display, which has the larger mean and which has the larger standard deviation?**

(a) Larger mean, A; larger standard deviation, A
(b) Larger mean, A; larger standard deviation, B
(c) Larger mean, B; larger standard deviation, A
(d) Larger mean, B; larger standard deviation, B
(e) Larger mean, B; same standard deviation



2. **The average cost per ounce for glass cleaner is 7.7 cents with a standard deviation of 2.5 cents. What is the z-score of the glass cleaner, Windex, that costs 10.1 cents per ounce?**

(a) 0.96 (b) 1.31 (c) 1.94 (d) 2.25 (e) 3.00

$z = \frac{10.1 - 7.7}{2.5}$

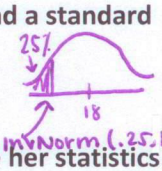
3. **What characteristic of a distribution does standard deviation measure?**

(a) shape (b) center (c) spread (d) skewness (e) frequency

4. **Scores on the American College Test (ACT) are normally distributed with a mean of 18 and a standard deviation of 6. The interquartile range of the scores is approximately:**

(a) 8.1 (b) 12 (c) 6 (d) 10.3 (e) 7

IQR = Q3 − Q1
75% − 25%



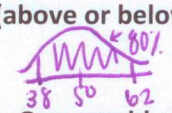
invNorm(.25, 18, 6)

5. **Ms. Jackson's Algebra II class had a standard deviation of 2.4 on their last test test, while her statistics class had a standard deviation of 1.2 on their last test. What can be said about these two classes? (The word homogeneous means alike, consistent, similar)**

(a) The algebra class's scores are more homogeneous than the statistics class's scores.
(b) The statistics class's scores are more homogeneous than the algebra class's scores.
(c) The statistics class did less well on the test than the algebra class.
(d) The algebra class performed twice as well on their test as did the statistics class.
(e) The algebra class performed 1.2 points better on their test than did the statistics class.

6 **The test grades at a large school have an approximately normal distribution with a mean of 50. What is the standard deviation of the data so that 80% of the students are within 12 points (above or below) the mean?**

(a) 5.875 (b) 9.375 (c) 10.375 (d) 14.5



$1.28 = \frac{62 - 50}{\sigma}$

7. **In a frequency distribution of 3000 scores, the mean is 78 and the median as 95. One would expect this distribution to be:**

(a) skewed to the right (b) skewed to the left (c) bimodal (d) symmetrical

8. **The stemplot displays the 1988 per capita income (in hundreds of dollars) of the 50 states. Which of the following best describes the data?**

(a) Skewed distribution, mean greater than median
(b) Skewed distribution, median greater than mean
(c) Symmetric distribution, mean greater than median
(d) Symmetric distribution, median greater than mean
(e) Symmetric distribution with outliers on high end

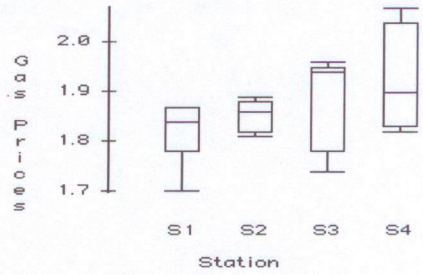
```
11|07
12|02255677788
13|23456777789
14|12667889
15|0002244559
16|6889
17|667
18|9
19|8
20|8
```

9. **Which of the following are true statements?**

✓ I. The standard deviation is the square root of the variance.
✓ II. The standard deviation is zero only when all values are the same.
✓ III. The standard deviation is strongly affected by outliers.

(a) I and II (b) I and III (c) II and III (d) I, II, and III (e) I only (f) III only

10. **A resident of Auto Town was interested in finding the cheapest gas prices at nearby gas stations. On randomly selected days over a period of one month, he recorded the gas prices (in dollars) at four gas stations near his house. The box plots of gas prices are as follows: Which station has more consistent gas prices?**



(a) Station 1 (b) Station 2 (c) Station 3
(d) Station 4 (e) Cannot be determined

11. **A small kiosk at the Atlanta airport carries souvenirs in the price range of $3.99 to $29.99, with a mean price of $14.75. The airport authorities decide to increase the rent charged for a kiosk by 5 percent. To make up for the increased rent, the kiosk owner decides to increase the prices of all items by 50 cents. As a result, which of the following will happen?**

mean ↑ std dev. =
range =

(a) The mean price and the range of prices will increase by 50 cents.
(b) The mean price will remain the same, but the range of prices will increase by 50 cents.
(c) The mean price and the standard deviation of prices will increase by 50 cents.
(d) The mean price will increase by 50 cents, but the standard deviation of prices will remain the same.
(e) The mean price and the standard deviation of prices will stay the same.

12. **The weights of cockroaches living in a typical college dormitory are approximately normally distributed with a mean of 80 grams and a standard deviation of 4 grams. The percentage of cockroaches weighing between 77 grams and 83 grams is about:**

(a) 99.7% (b) 95% (c) 68% (d) 55% (e) 34%

normalcdf(77,83,80,4)

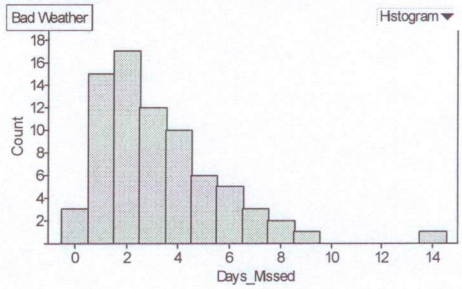
13. **Which of the following are true statements?**

✓ I. In all normal distributions, the mean and median are equal.
✓ II. All bell-shaped curves are normal distributions for some value of μ and σ.
✓ III. Virtually all the area under a normal curve is within three standard deviations of the mean, no matter what the particular mean and standard deviation are.

(a) I and II (b) I and III (c) II and III (d) I, II, and III (e) I only

14. **In the northern U.S., schools are sometimes closed during winter due to severe snowstorms. At the end of the school year, schools have to make up for the days missed. The following graph shows the frequency distribution of the number of days missed due to snowstorms per year, using data from the past 75 years. Which of the following should be used to describe the center of the distribution?**

(a) Mean, because it is an unbiased estimator.
(b) Median, because the distribution is skewed.
(c) IQR, because it excludes outliers and includes only the middle 50 percent of the data.
(d) First quartile, because the distribution is left skewed.
(e) Standard deviation, because it is unaffected by outliers.



15. **A large company has offices in two locations, one in New Jersey and one in Utah. The mean salary of the office assistants in the New Jersey office is $28,500. The mean salary of office assistants in the Utah office is $22,500. The New Jersey office has 128 office assistants and the Utah office has 32 office assistants. What is the mean salary paid to the office assistants in this company?**

(a) $22,500 (b) $23,700 (c) $25,500 (d) $27,300 (e) $28,500

$\frac{128}{160} = .8$ $\frac{32}{160} = .2$ $.8(28{,}500) + .2(22{,}500)$

16. **A distribution of 6 scores has a median of 21. If the highest score increase 3 points, what will be the value of the median?**

(a) 21 (b) 21.5 (c) 24 (d) 27 (e) cannot be determined with the information given

Rainwater was collected in water collectors at thirty-one different sites near an industrial basin and the amount of acidity (pH level) was measured. The following stem plot shows the pH values that ranged from 2.6 to 6.3.

Rainwater pH

2	679
3	237789
4	1222446899
5	05567888
6	0233

4|2 = 4.2 pH

17. **What is the median pH reading?**

(a) 4.2 (b) 4.4 (c) 4.5 (d) 4.6

18. **Which boxplot represents the data in the stemplot?**

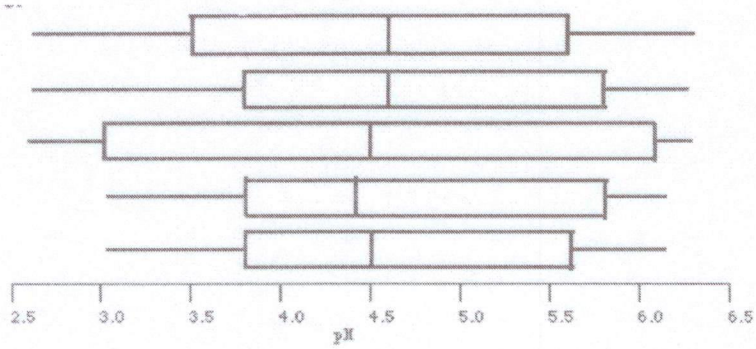
(a)

(b)

(c)

(d)

(e)



19. **What is the interquartile range?**

(a) 1.9 (b) 3.7 (c) 3.8 (d) 4.5 (e) 5.6

5.7 − 3.8 = 1.9

20. **The following is a stem plot of the birth weights of 26 male babies born to a smoking group of mothers. What is the third quartile weight, in kilograms, of the male babies in this sample?**

(a) 42 (b) 4.3 (c) 3.5 (d) 4.2

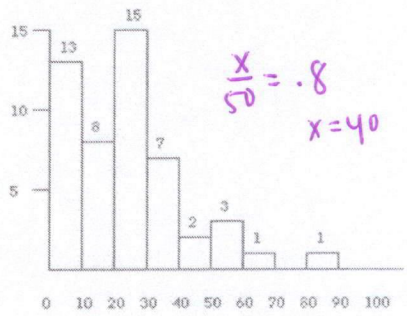
Birth Weight of Male Babies

2	346778889
3	22346789
4	1224
5	3556

2|3 = 2.3 kg

21. **The following is a histogram showing the actual frequency of the closing prices for 50 days of trading on the New York exchange for stock XYZ. Which class contains the 80th percentile?**

(a) 10 – 20 (b) 20 – 30 (c) 30 – 40 (d) 40 – 50 (e) 50 – 60



$\frac{x}{50} = .8$

$x = 40$

22. **In the histogram, what best describes the shape of the distribution?**

(a) symmetric (b) skewed left

(c) skewed right (d) skewed in both directions

23. **Use the following output from the statistical software Data Desk when analyzing the pH values of data collected on precipitation events in 1986. Which of the following is <u>not correct</u>?**

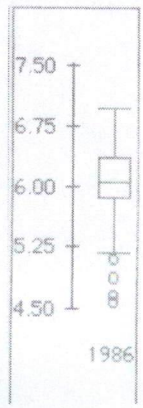
(a) The interquartile range is about 0.34

(b) The 25th percentile is about 5.9.

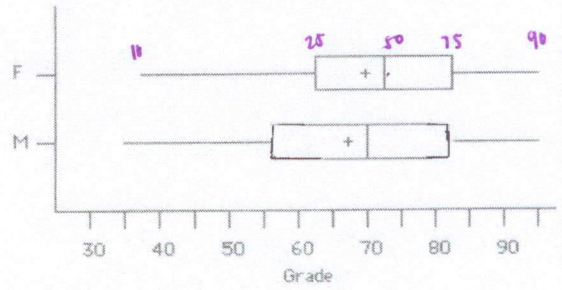
(c) The median is about 5.24.

(d) About 75% of the data is less than 6.4.

(e) Some outliers appear to be present below a pH of 5.25.



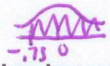
24. **Consider the following box plots of males (M) and females (F) for grades in a course in statistics. These boxplots are drawn according to the convention that the whiskers only reach to the 10th and 90th percentiles, not the minimum and maximum values. The "+" indicates the location of the mean. Which of the following is correct?**



(a) The mean grade of the female students is about 72.
(b) The median of the male students is about 60.
(c) The male IQR has more variability than the female IQR.
(d) About 25% of the female students get grades above 72.
(e) About 10% of the male students get grades below 56.

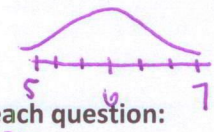
25. **Which of these variables is least likely to have a Normal distribution?**
(a) Annual income for all 150 employees at a local high school — skewed because teachers and administrators are paid differently
(b) Lengths of 50 newly hatched pythons
(c) Heights of 100 white pine trees in a forest
(d) Amount of soda in 60 cups filled by an automated machine at a fast-food restaurant
(e) Weights of 200 of the same candy bar in a shipment to a local supermarket

26. **The proportion of observations from a standard Normal curve that take values larger than -0.75 is about**
(a) 0.2266 (b) 0.7704 (c) .7734 (d) 0.7764 (e) .8023



normalcdf(-.75, ∞, 0, 1)

27. **If the heights of American men follow a Normal distribution, and 99.7% have heights between 5'0" and 7'0", what is the best estimate of the standard deviation of the heights of American men?** 3 standard deviations
(a) 1" (b) 3" (c) 4" (d) 6" (e) 12"



28. **Without using the lists in your calculator, use the following set of data to answer each question:**

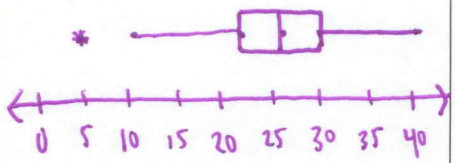
5	11	14	17	21	22 23	23	24	25	26	26	28	29	29	30	30	31	35	35	41

Check for outliers using the formula

IQR = 30 − 22 = 8
1.5(IQR) = 1.5(8) = 12
Q3 + 12 = 30 + 12 = 42
Q1 − 12 = 22 − 12 = 10

*5 is a low outlier
*no high outliers

Make a boxplot, by hand



Make a stemplot, by hand

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

1|4 = 14

don't forget!

29. **Use your formula packet, if necessary, to calculate the following probabilities. Given that $P(A) = 0.35$, $P(B) = 0.6$, and $P(A \cap B) = 0.27$:**

$P(A^c) =$ 1 − .35 = .65

$P(A \cup B) =$ .68

$P(A|B) =$ .27 / .6 = .45

Use **both** formulas of independence to show that these events are not independent:

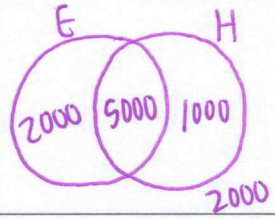
$P(A \cap B) = P(A) \cdot P(B)$ → $.27 \neq .35(.6)$

$P(A|B) = P(A)$ → $.45 \neq .35$

REVIEW OF PROBABILITY

30. Of the 10,000 freshman at BYU, 7000 must take English, 6000 must take History, and 5000 must take both. Suppose that a student is randomly selected.

a) What is the probability that the selected student must take only English?



$\frac{2000}{10000} = .2$

b) What is the probability that the selected student must take neither English nor History?

$\frac{2000}{10,000} = .2$

c) Suppose you learn that the selected student must take English, what is the probability that this student must take both English and History?

$P(\text{Both} \mid \text{English}) = \frac{5000}{7000} = .714$

d) Are the outcomes *must take English and must take History* independent? Explain.

$P(E \text{ and } H) \stackrel{?}{=} P(E) \cdot P(H)$

$.5 = .7 \cdot .6$

NO

$P(E \mid H) \stackrel{?}{=} P(E)$

$\frac{.5}{.6} = .7$

NO

REVIEW OF BINOMIAL

31. The AP Statistics exam includes 40 multiple choice questions, each with 5 answer choices. Suppose you have forgotten everything and must guess on every question. Let *x* represent the number of correct responses.

a) What kind of ~~probability distribution~~ does *x* have? Explain. Binomial

Bi - right or wrong
nom - 40 = n
i - each question independent
al - p = .2

b) What is your expected score on the exam?

$\mu_x = np =$

$40(.2) = 8$

c) Compute the variance and standard deviation of *x*?

$\sigma_x = \sqrt{npq} = 2.53$ ← std. dev

$\sigma_x^2 = npq = 6.4$ ← variance

d) What is the probability that you will get exactly 25 questions correct?

binompdf (40, .2, 25) = .00000005 ≈ 0

(n, p, x)

e) What is the probability that you correctly answer at least 16 problems?

1 - binomcdf (40, .2, 15) = .0029

(n, p, x)

REVIEW OF SAMP. DISTRIBUTIONS

31. A plane used to fly tourists in and out of the rain forest contains seating for 16 passengers. The total weight limit for the passengers is 2500 pounds. Assume the average weight of tourists is 150 pounds, the standard deviation 27 pounds, and that the distribution of tourist weights is approximately normal. If the weight limit is exceeded, the plane has difficulty taking off safely. If a random sample of 16 tourists has booked a flight, what is the chance that the weight limit will be exceeded?

$\mu_{\bar{x}} = \mu = 150$

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{27}{\sqrt{16}} = 6.75$

normalcdf (156.25, ∞, 150, 6.75) = .177

OR

normalcdf (2500, ∞, 2400, 108) = .177

32. A manufacturer of iPods purchases computer chips from a vendor. When a large shipment is received, a random sample of 200 computer chips is selected, and each is inspected. If the sample proportion of defectives is more than .02, the entire shipment will be returned to the vendor. What is the approximate probability that the shipment will be returned if the true proportion of defectives in the shipment is .05?

$\mu_{\hat{p}} = p = .05$

$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(.05)(.95)}{200}} = .015$

normalcdf (.02, ∞, .05, .015) = .977

2004 #2 Free Response

Researchers who are studying a new shampoo formula plan to compare the condition of hair for people who use the new formula with the condition of hair for people who use the current formula. Twelve volunteers are available to participate in this study. Information on these volunteers (numbered 1 through 12) is shown in the table below.

Volunteer	Gender	Age
1	Male	21
2	Female	20
3	Male	47
4	Female	60
5	Female	62
6	Male	61
7	Male	58
8	Female	44
9	Male	44
10	Female	24
11	Male	23
12	Female	46

a) These researchers want to conduct an experiment involving the two formulas (new and current) of shampoo. They believe that the condition of hair changes with age, but not gender. Because researchers want the size of the blocks in an experiment to be equal to the number of treatments, they will use blocks of size 2 in their experiment. Identify the volunteers (by number) that would be included in each of the six blocks and give the criteria you used to form the blocks.

Blocks: (1,2) (10,11) (8,9) (3,12) (4,7) (5,6)

The two youngest were paired, the next two youngest, etc. until all 12 are part of a block.

b) Other researchers believe that hair condition differs with both age and gender. These researchers will also use blocks of size 2 in their experiment. Identify the volunteers (by number) that would be included in each of the six blocks and give the criteria you used to form the blocks.

Blocks (2,10) (8,12) (4,5) (1,11) (3,9) (6,7)

The two youngest women were paired, the next two youngest, and the two oldest. same with men.

c) The researchers in part (b) decide to select three of the six blocks to receive the new formula and to give the other three blocks the current formula. Is this an appropriate way to assign treatments? If so, describe a method for selecting the three blocks to receive the new formula. If not, describe an appropriate method for assigning treatments.

No, they blocked on age and gender because they believe those are confounding variables. They should randomly select one person from each block to receive the new formula and the other the current one.