Instructor: M. Pivovarova

Duration: 55 minutes: 50 minutes are for writing the quiz and 5 minutes are for the TA's and me to distribute and collect papers.

Format: 20 multiple-choice questions. Record answers on pink SCANTRON form. Correct answers are worth 5 points each and incorrect answers are worth 0 points. There are 100 total possible points.

Allowed aids: A non-programmable calculator and aid sheets provided.

Instructions:

- Answers must be properly recorded on the pink SCANTRON form to earn marks
- Print your **LAST NAME** and **INITIALS** in the boxes AND darken each letter in the corresponding bracket below each box; Sign your name in the **SIGNATURE** box
- Print your 9 digit **STUDENT NUMBER** in the boxes AND darken each number in the corresponding bracket below each box
- Your **FORM NUMBER** is **01**
- Use only a pencil or blue or black ball point pen
    - o **Pencil strongly recommended** because it can be erased
- Make dark solid marks that fill the bubble completely
- Erase completely any marks you want to change
- Crossing out a marked box is not acceptable and is incorrect
- If more than one answer is marked then that question earns 0 points

(1) Fill in the missing words to the quote: "Statistical methods may be described as methods for drawing conclusions about _____ based on _____ computed from the _____ ."

>   **(A)** statistics, samples, populations
>   **(B)** populations, parameters, samples
>   **(C)** populations, statistics, samples
>   **(D)** parameters, statistics, populations
>   **(E)** statistics, parameters, samples

(2) Since the population size is always larger than the sample size, then the sample statistic:

>   **(A)** Can never be larger than the population parameter
>   **(B)** Can never be equal to the population parameter
>   **(C)** Can never be zero
>   **(D)** Can never be smaller than the population parameter
>   **(E)** None of the above answers is correct

► <u>For Question **(3)**</u>, consider Figure 1.
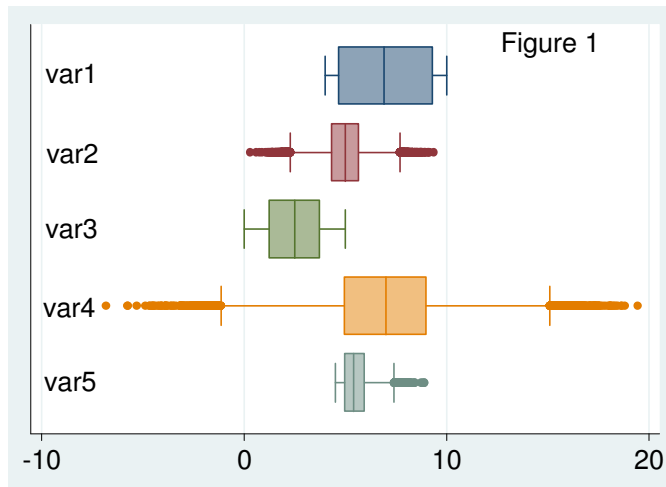
Figure 1

(3) Which of the variables summarized in Figure 1 could be normally distributed (assuming large sample size)?

>   **(A)** var1 only
>   **(B)** var5 only
>   **(C)** var1 and var3
>   **(D)** var2 and var4
>   **(E)** var2, var3 and var4

(4) Molly earned a score of 940 on a national achievement test. The mean test score was 850 with a standard deviation of 100. Which statement characterizes Molly's relative standing among the students who took the test? (Assume that test scores are normally distributed.)

        (A) Approximately 32% of students did better than Molly
        (B) At least 75% of students did worse than Molly
        (C) At least 88.9% did worse than Molly
        (D) Approximately 16% of students did better than Molly
        (E) Not enough information

(5) A randomized experiment was done by randomly assigning each participant either to walk for half an hour three times a week or to sit quietly reading a book for half an hour three times a week. At the end of a year the change in participants' blood pressure over the year was measured, and the change was compared for the two groups. This is a randomized experiment rather than an observational study because:

        (A) Blood pressure was measured at the beginning and end of the study.
        (B) The two groups were compared at the end of the study.
        (C) The participants were randomly assigned to either walk or read, rather than choosing their own activity.
        (D) The participants were compared at the beginning of the study.
        (E) A random sample of participants was used.

(6) A national consumer magazine reported the following correlations. The correlation between car weight and car reliability is -0.30. The correlation between car weight and annual maintenance cost is 0.20. Which of the following statements are true?

I. Heavier cars tend to be less reliable.
II. Heavier cars tend to cost more to maintain.
III. Car weight is related more strongly to reliability than to maintenance cost.

        (A) I only
        (B) II only
        (C) III only
        (D) I and II
        (E) I, II, and III

(7) A survey of rural farmers in Burkina Faso had been undertaken in order to measure the impact of fertilizer on cotton yield. Each farmer was asked about average amount of fertilizer he/she used and average yield per plot. These are _____ data that contain _____ variables.

(A) Panel; interval
(B) Cross-sectional; interval
(C) Time-series; interval
(D) Cross-sectional; nominal
(E) Panel; nominal

(8) Researcher is interested whether individuals with different lifestyles also differ in their dieting habits. The data were collected for 489 individuals on the number of meals they consume per day. The mean of the sample is 2.5 meals per day and standard deviation is equal to 2. What is the most reasonable statement about the shape of the distribution:

(A) It is symmetric
(B) It is bimodal
(C) It is negatively skewed
(D) It is bell-shaped
(E) It is positively skewed

(9) In a sample of 90 customers who were asked how many times per month they do grocery shopping, the following descriptive statistics have been obtained: mean = 9 and standard deviation = 3. The histogram is bell-shaped. Which statement about the distribution is correct?

(A) 62 customers do grocery shopping between 6 and 12 times per month
(B) 62 customers do grocery shopping between 3 and 15 times per month
(C) 68 customers do grocery shopping between 6 and 12 times per month
(D) 86 customers do grocery shopping between 6 and 12 times per month
(E) 80 customers do grocery shopping between 3 and 15 times per month

(10) Consider the descriptive statistics for the total sales of the product and advertising budget (measured in thousands of dollars). Coefficient of correlation between sales and advertising is 0.65.

	Sales	Advertising
Mean	152	11
St. Dev.	30	5

What is the predicted value of sales if advertising budget is equal to 10 (round to the nearest decimal point).

(A) 151.9
(B) 151.4
(C) 148.1

    (D) 142.0
    (E) 128.1

(11) Which of the following conditions is necessary for the population mean to be equal to the population median?

    (A) The population must be normal (bell-shaped)
    (B) The population must be unimodal
    (C) The population must be symmetric
    (D) The population must have variance greater than zero
    (E) All of the above

(12) The distribution of a random variable W is bell-shaped. If 95% of observations are between 75 and 115, what is the mean of the distribution?

    (A) 80
    (B) 90
    (C) 95
    (D) 105
    (E) 110

►For Questions **(13)-(14),** consider these summary statistics for a random sample:

$$\bar{X}=5 \text{ , } \bar{Y}=15 \text{ , } s_x=2, \ s_y=5, \ s_{xy}=-8$$

(13) For which line would the SSE (sum of the squared errors) be <u>smallest</u>?

    (A) Y-hat = 5 + 2*X
    (B) Y-hat =25 − 2*X
    (C) Y-hat =33 - 1.6*X
    (D) Y-hat = -5 + 4*X
    (E) Y-hat =35 − 4*X

(14) What is the coefficient of determination?

    (A) 0.64
    (B) -0.80
    (C) 1.20
    (D) -4.00
    (E) 64.0

► For Questions **(15)-(16)**, consider this histogram of a random sample of variable X.

n = 125
mean = 3.36, sd = 1.96
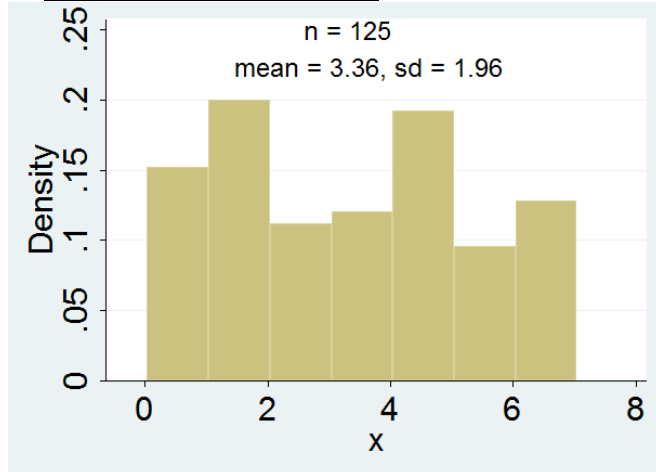
(15) About what percent of observations are between 1 and 2?

        (A) 12%
        (B) 16%
        (C) 20%
        (D) 24%
        (E) 30%

(16) Given this sample, what is a reasonable inference about the <u>population</u> median?

        (A) It is equal to the population mean
        (B) It is equal to 3.36
        (C) It is greater than 3.36
        (D) It is less than 3.36
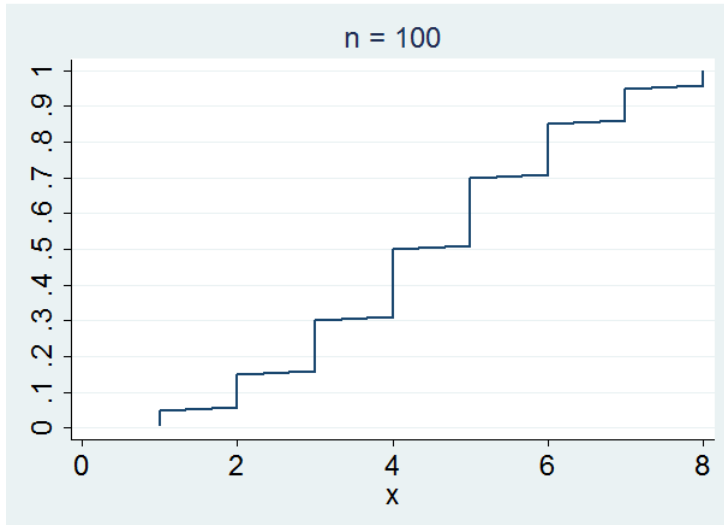   **(E)** It is approximately 3

(17) Consider tabulation of a random variable Y. What is the coefficient of variation? (Round to the second decimal point)

X	Freq.	Percent	Cum.
0.1	15	30.00	30.00
0.2	20	40.00	70.00
0.3	5	10.00	80.00
0.4	10	20.00	100.00
Total	50	100.00	

        (A) 0.01
        (B) 0.05
        (C) 0.11

(D) 0.22
(E) 0.50

► For Question **(18)**, consider ogive of a random sample of variable X.

n = 100

(18) Approximately how many observations are greater than 6?

(A) 10
(B) 15
(C) 20
(D) 25
(E) 30

► For Questions **(19)-(20)**, consider joint probability table based on the results from the customer satisfaction survey in a restaurant chain:

Rating	Customer Will Return	Customer Will Not Return
Poor	.02	.10
Fair	.08	.09
Good	.35	.14
Excellent	.20	.02

(19) Rating (event A) and decision to return (event B) are independent events because:

(A) Conditional probability of A is not equal to conditional probability of B
(B) Marginal probability of A is not equal to conditional probability of A
(C) Marginal probability of A is equal to conditional probability of A

(D) Events A and B are not independent
(E) We can use multiplication rule for independent events

(20) What is the chance that a customer who rated the restaurant as "poor" will return to that restaurant?

(A) .02
(B) .10
(C) .12
(D) .17
(E) .65

Make sure to properly mark your name, student number, signature, and answers for Part 1 on the pink SCANTRON form <u>before</u> the end of the quiz is announced. No extra time is permitted.

Your form number is 01. Complete the FORM box at the top right of your pink SCANTRON form.