Student: ____________________ **Date:** ____________________	**Instructor:** Richeng Piao **Course:** ECON 2560 - Applied Econometrics	**Assignment:** Practice Problem Set 2

1. Determine whether the following variables are random or not:

   1. The time it takes to commute to school. (1) __________

   2. The gender of the next person you will meet. (2) __________

   3. The daily return of a stock. (3) __________

   4. The number of days in a week. (4) __________

   5. The number of times it rains during the summer. (5) __________

   (1) ○ Random ○ Not random
   (2) ○ Random ○ Not random
   (3) ○ Random ○ Not random
   (4) ○ Not random ○ Random
   (5) ○ Random ○ Not random

   Answers (1) Random

   (2) Random

   (3) Random

   (4) Not random

   (5) Random

   ID: Review Concept 2.1

---

2. Suppose that the random variables X and Y are independent and you know their distributions.

   Which of the following explains why knowing the value of X tells you nothing about the value of Y?

   ○ **A.** The variance of X might be different from the variance of Y.
   ○ **B.** The mean of X might be different from the mean of Y.
   ○ **C.** X and Y might be independent.
   ○ **D.** All of the above.

   Answer: C. X and Y might be independent.

   ID: Review Concept 2.2

---

3. Suppose that X denotes the amount of rainfall in your hometown during a given month and Y denotes the number of children born in Los Angeles during the same month.

Which of the following statements best explains why X and Y are not independent?

- **A.** The variance for the amount of rainfall in inches is equal to the variance of the number of children born.
- **B.** The ratio of the amount of rainfall in inches to the number of children born is usually one.
- **C.** The amount of rainfall may tell you something about the season and, since births are seasonal, it may also tell you something about the number of children born.
- **D.** The expected value of rainfall in inches is equal to the expected number of children born.

Answer: C.

The amount of rainfall may tell you something about the season and, since births are seasonal, it may also tell you something about the number of children born.

ID: Review Concept 2.3

---

4. Let X be the number of applicants who apply for a senior level position at a large multinational corporation. The probability distribution of the random variable X is given in the following table. The outcomes (number of applicants) are mutually exclusive.

Complete the table by calculating the cumulative probability distribution of X.

	Outcome (Number of applicants)				
	0	**1**	**2**	**3**	**4**
Probability distribution	0.40	0.25	0.15	0.15	0.05
Cumulative probability distribution					

The probability that there will be at least two applicants is ____, and the probability that there will be at most three applicants is ____.

The probability that there will be three or four applicants is ____.

Answers 0.40

0.65

0.80

0.95

1.00

0.35

0.95

0.20

ID: Concept Exercise 2.1.1

---

5. An econometrics class has 80 students, and the mean student weight is 145 lb. A random sample of four students is selected from the class, and their average weight is calculated. Will the average weight of the students in the sample equal 145lb?

- **A.** Yes.
- **B.** No.

Using this example, which of the following best explains the sample average $\bar{Y}$?

- **A.** The value of $\bar{Y}$ is random. However, it is the same for all samples.
- **B.** Because each observation Y_i is drawn at random, the value of their average, $\bar{Y}$, is also random. The value of $\bar{Y}$ differs from one sample to the next.
- **C.** Although each observation Y_i is random, the value of their average, $\bar{Y}$, is not random. $\bar{Y}$ is equal to the population mean.
- **D.** The value of $\bar{Y}$ is not random, but it differs from one sample to the next.

Answers B. No.

B.
Because each observation Y_i is drawn at random, the value of their average, $\bar{Y}$, is also random. The value of $\bar{Y}$ differs from one sample to the next.

ID: Review Concept 2.4

6. Observe that for a random variable Y that takes on values 0 and 1, the expected value of Y is defined as follows:

$$E(Y) = 0 \times \Pr(Y=0) + 1 \times \Pr(Y=1)$$

Now, suppose that X is a Bernoulli random variable with success probability Pr $(X = 1) = p$. Use the information above to answer the following questions.

Show that $E\left(X^4\right) = p$.

$$E\left(X^4\right) = (\square \times \square) + (\square \times p) = \square$$

(*Use the tool palette on the right to insert superscripts. Enter you answer in the same format as above.*)

Suppose that $p = 0.28$.

Compute the mean of X.

$E(X) =$ $\square$

(*Round your response to two decimal places*)

Compute the variance of X.

$var(X) =$ $\square$

(*Round your response to three decimal places*)

Compute the skewness of X using the following formula:

$$\frac{E(X-E(X))^3}{\sigma^3} = \frac{E\left(X^3\right) - 3\left[E\left(X^2\right)\right][E(X)] + 2[E(X)]^3}{\sigma^3}$$

Skewness of X = $\square$

(*Round your response to three decimal places*)

Compute the kurtosis of X using the following formula:

$$\frac{E(X-E(X))^4}{\sigma^4} = \frac{E\left(X^4\right) - 4[E(X)]\left[E\left(X^3\right)\right] + 6[E(X)]^2\left[E\left(X^2\right)\right] - 3[E(X)]^4}{\sigma^4}$$

Kurtosis of X = $\square$

(*Round your response to three decimal places*)

Answers 0

1 – p

1

p

0.28

0.202

0.980

1.960

ID: Exercise 2.4

7. In September, Seattle's daily high temperature has a mean of 70°F and a standard deviation of 13°F.

The formula to convert degrees Fahrenheit °F to degrees Celsius °C is:

$$^{\circ}C = \frac{5}{9}\left(^{\circ}F - 32\right)$$

Use this information to answer the following questions.

Compute the mean of Seattle's daily high temperature in degrees Celsius °C.

The mean of the daily high temperature in degrees Celcius = ☐ °C

(*Round your response to three decimal place.*)

Compute the standard deviation of Seattle's daily high temperature in degrees Celsius °C.

The standard deviation of the daily high temperature in degrees Celisius = ☐ °C

(*Round your response to three decimal places*)

Compute the variance of Seattle's daily high temperature in degrees Celsius °C.

The variance of the daily high temperature in degrees Celsius = ☐ °C

(*Round your response to three decimal places*)

Answers 21.111

7.222

52.157

ID: Exercise 2.5

---

8. Suppose the random variable Y has a mean of 47 and a variance of 36. Let $Z = \frac{1}{\sqrt{36}}(Y - 47)$.

Show that $\mu_Z = 0$.

$$\mu_Z = E\left[\boxed{}\left(Y - \boxed{}\right)\right]$$
$$= \boxed{}\left[\mu_Y - \boxed{}\right] = 0$$

(*Round your responses to two decimal places*)

Show that $\sigma_Z^2 = 1$.

$$\sigma_Z^2 = var\left[\boxed{}\left(Y - \boxed{}\right)\right]$$
$$= \boxed{}\sigma_Y^2 = 1$$

(*Round your responses to two decimal places*)

Answers 0.17

47

0.17

47

0.17

47

0.03

ID: Exercise 2.8

---

9. The expected value of a discrete random variable:

- **A.** equals the population median.
- **B.** is computed as a weighted average of the possible outcome of that random variable, where the weights are the probabilities of that outcome.
- **C.** is the outcome that is most likely to occur.
- **D.** can be found by determining the 50% value in the c.d.f.

Answer: B.

is computed as a weighted average of the possible outcome of that random variable, where the weights are the probabilities of that outcome.

ID: Test A Ex 2.2.1

---

10. The variance of $\overline{Y}, \sigma^2_{\overline{Y}}$ is given by the following formula:

- **A.** $\frac{\sigma^2_Y}{\sqrt{n}}$.
- **B.** σ^2_Y.
- **C.** $\frac{\sigma_Y}{\sqrt{n}}$.
- **D.** $\frac{\sigma^2_Y}{n}$.

Answer: D. $\frac{\sigma^2_Y}{n}$.

ID: Test A Ex 2.2.2

---

11. $\sum_{i=1}^{n} \left(ax_i + b\right) =$

- **A.** $n \times a \times \bar{x} + n \times b$.
- **B.** $\bar{x} + n \times b$.
- **C.** $n(a + b)$.
- **D.** $n \times a + \bar{x}$.

Answer: A. $n \times a \times \bar{x} + n \times b$.

ID: Test A Ex 2.2.3

---

12. The mean and variance of a Bernoulli random variable are given as:

- **A.** p and $p(1 - p)$.
- **B.** np and $np(1 - p)$.
- **C.** p and $\sqrt{p(1 - p)}$.
- **D.** cannot be calculated.

Answer: A. p and $p(1 - p)$.

ID: Test A Ex 2.2.4

---

13. For a normal distribution, the *skewness* and *kurtosis* measures are as follows:

- **A.** 1.96 and 4.
- **B.** 0 and 0.
- **C.** 0 and 3.
- **D.** 1 and 2.

Answer: C. 0 and 3.

ID: Test B Ex 2.2.1

---

14. $\sum_{i=1}^{n} (ax_i + by_i + c) =$

- **A.** $a \sum_{i=1}^{n} x_i + b \sum_{i=1}^{n} y_i + c.$
- **B.** $a \sum_{i=1}^{n} x_i + b \sum_{i=1}^{n} y_i.$
- **C.** $a \sum_{i=1}^{n} x_i + b \sum_{i=1}^{n} y_i + n \times c.$
- **D.** $a\bar{x} + b\bar{y} + n \times c.$

Answer: C. $a \sum_{i=1}^{n} x_i + b \sum_{i=1}^{n} y_i + n \times c.$

ID: Test B Ex 2.2.2

---

15. Assume that you assign the following subjective probabilities for your final grade in your econometrics course (the standard GPA scale of 4 = A to 0 = F applies):

Grade	Probability
A	0.20
B	0.50
C	0.20
D	0.08
F	0.02

The expected value is:

- **A.** 3.25.
- **B.** 3.0.
- **C.** 2.78.
- **D.** 3.5.

Answer: C. 2.78.

ID: Test B Ex 2.2.3

---

16. Consider the following linear transformation of a random variable

$$y = \frac{x - \mu_x}{\sigma_x},$$

where μ_x is the mean of x and σ_x is the standard deviation. Then the expected value and the standard deviation of Y are given as:

- **A.** 0 and 1.
- **B.** 1 and 1.
- **C.** $\frac{\mu_x}{\sigma_x}$ and σ_x.
- **D.** cannot be computed because Y is not a linear function of X.

Answer: A. 0 and 1.

ID: Test B Ex 2.2.4

---

17. In a basketball match, a shooting guard can either throw the ball in the basket and score points for his team or miss it and not score. Each successful basket is worth one point.

The shooting guard's ball throw is an example of a (1) ____________ random variable, because the outcomes are (2) ____________.

Suppose the shooting guard successfully throws the ball in the basket 3 out of 10 times. Complete the probability distribution of the shooting guard's ball throws (X).

X	Successful throw	Unsuccessful throw
$P(X)$		

The expected value of the shooting guard's throw is ______, and its variance is ______.

Now assume that each successful basket is worth two points.

The expected value of the shooting guard's throw is ______ and its variance is ______.

(1) ○ continuous ○ Bernoulli

(2) ○ a continuum of possible values ○ binary

Answers (1) Bernoulli

(2) binary

0.3

0.7

0.3

0.21

0.6

0.84

ID: Concept Exercise 2.2.1

---

18. Suppose Walmart introduces an offer of a flat 20% discount on the entire bill for the purchase of any electronic item. It also offers an additional $100 discount on the entire bill for the purchase of any kitchenware item, conditional on the purchase of an electronic item. Let X denote the before-discount expenditure of a shopper who purchases both an electronics item and a kitchenware item, and let Y denote this shopper's after-discount expenditure.

If σ_X denotes the standard deviation of this shopper's before-discount expenditure, the standard deviation of the after-discount expenditure of this shopper is ______.

(*Carefully enter your response as an algebraic expression, using the proper notation in the proper format.*)

The value of variance is

- **A.** always non-negative.
- **B.** always negative.
- **C.** always positive.
- **D.** none of the above.

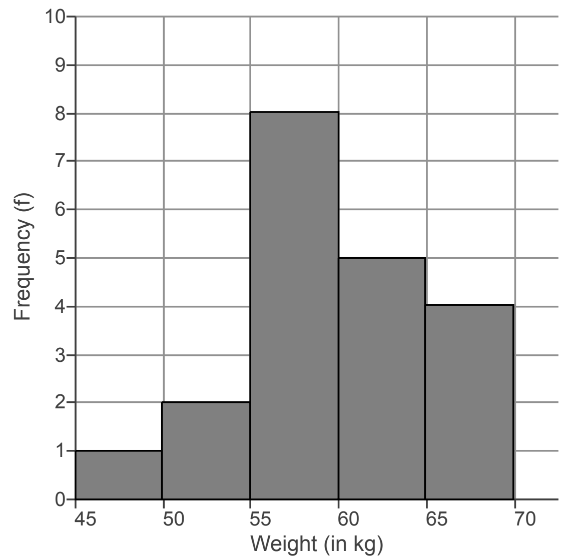
Answers $0.80\sigma_X$

A. always non-negative.

ID: Concept Exercise 2.2.2

19. Which of the following expressions is used to calculate the skewness of a distribution?

- A. $\dfrac{E\left[\left(Y-\mu_Y\right)^3\right]}{\sigma_Y^4}$.
- B. $\dfrac{E\left[\left(Y-\mu_Y\right)^4\right]}{\sigma_Y^4}$.
- C. $\dfrac{E\left[\left(Y-\mu_Y\right)^2\right]}{\sigma_Y^2}$.
- D. $\dfrac{E\left[\left(Y-\mu_Y\right)^3\right]}{\sigma_Y^3}$.



The given figure shows a histogram. The histogram suggests that the distribution is (1) __________.

The kurtosis of this distribution is (2) __________.

(1)
- symmetrical
- negatively skewed
- positively skewed

(2)
- negative
- positive

Answers D. $\dfrac{E\left[\left(Y-\mu_Y\right)^3\right]}{\sigma_Y^3}$.

(1) negatively skewed

(2) positive

ID: Concept Exercise 2.2.3

20. Use the probability distribution given in the table below to answer the following questions

Joint Distribution of Weather Conditions and Commuting Times

	Rain ($X = 0$)	No Rain ($X = 1$)	**Total**
Long commute ($Y = 0$)	0.07	0.08	0.15
Short commute ($Y = 1$)	0.82	0.03	0.85
Total	0.89	0.11	1.00

Compute the mean of Y.

$E(Y) =$ ______

(*Round your response to two decimal places*)

Compute the mean of X.

$E(X) =$ ______

(*Round your response to two decimal places*)

Compute the variance of X.

$\sigma_X^2 =$ ______

(*Round your response to four decimal places*)

Compute the variance of Y.

$\sigma_Y^2 =$ ______

(*Round your response to four decimal places*)

Compute the covariance of X and Y.

$\sigma_{XY} =$ ______

(*Round your response to four decimal places*)

Compute the correlation of X and Y.

$corr(X, Y) =$ ______

(*Round your response to four decimal places*)

Answers 0.85

0.11

0.0979

0.1275

-0.0635

-0.5684

ID: Exercise 2.2

21. Use the probability distribution given in the table below and consider two new random variables, $W = 1 + 5X$ and $V = 9 + 4Y$, to answer the following questions

Joint Distribution of Weather Conditions and Commuting Times

	Rain ($X = 0$)	No Rain ($X = 1$)	**Total**
Long commute ($Y = 0$)	0.21	0.36	0.57
Short commute ($Y = 1$)	0.28	0.15	0.43
Total	0.49	0.51	1.00

Compute the mean of W.

$E(W) =$ ______

(*Round your response to two decimal places*)

Compute the mean of V.

$E(V) =$ ______

(*Round your response to two decimal places*)

Compute the variance of W.

$\sigma^2_W =$ ______

(*Round your response to four decimal places*)

Compute the variance of V.

$\sigma^2_V =$ ______

(*Round your response to four decimal places*)

Compute the covariance between W and V.

$\sigma_{WV} =$ ______

(*Round your response to four decimal places*)

Compute the correlation between W and V.

$corr(W, V) =$ ______

(*Round your response to four decimal places*)

Answers 3.55

10.72

6.2475

3.9216

– 1.3860

– 0.2800

ID: Exercise 2.3

22. The following table gives the joint probability distribution between employment status and college graduation among those either employed or looking for work (unemployed) in the working age U.S. population.

	Unemployed ($Y = 0$)	Employed ($Y = 1$)	Total
Non-college grads ($X = 0$)	0.0594	0.6389	0.6983
College grads ($X = 1$)	0.0145	0.2872	0.3017
Total	0.0739	0.926	0.9999

The expected value of Y, denoted $E(Y)$, is ______. (*Round your response to three decimal places.*)

The unemployment rate is the fraction of the labor force that is unemployed. Show that the unemployment rate is given by $1 - E(Y)$.

Unemployment rate $= 1 -$ ______ $= 1 - E(Y) = 1 - 0.926 = 0.0739$.

$E(Y \mid X = 1)$ is ______. (*Round your response to three decimal places.*)

$E(Y \mid X = 0)$ is ______. (*Round your response to three decimal places.*)

The unemployment rate for college graduates is ______, and the unemployment rate for non-college graduates is ______. (*Round your responses to three decimal places.*)

A randomly selected member of this population reports being unemployed. The probability that this worker is a college graduate is ______, and the probability that this worker is a non-college graduate is ______. (*Round your responses to three decimal places.*)

Are educational achievement and employment status independent?

- **A.** Since $\Pr(X = 0, Y = 1) = \Pr(X = 0)$, educational achievement and employment status are independent.
- **B.** Since $\Pr(X = 0 \mid Y = 1) = \Pr(X = 0)$, educational achievement and employment status are independent.
- **C.** Since $\Pr(X = 0, Y = 1) \neq \Pr(X = 0)$, educational achievement and employment status are not independent.
- **D.** Since $\Pr(X = 0 \mid Y = 1) \neq \Pr(X = 0)$, educational achievement and employment status are not independent.

Answers 0.926

$\Pr(Y = 1)$

0.952

0.915

0.048

0.085

0.196

0.804

D. Since $\Pr(X = 0 \mid Y = 1) \neq \Pr(X = 0)$, educational achievement and employment status are not independent.

ID: Exercise 2.6

23. Suppose you have some money to invest—for simplicity, $1—and you are planning to put a fraction w into a stock market mutual fund and the rest, $1 - w$, into a bond mutual fund. Suppose that $1 invested in a stock fund yields R_s after 1 year and that $1 invested in a bond fund yields R_b, suppose that R_s is random with mean 0.08 (8%) and standard deviation 0.07, and suppose that R_b is random with mean 0.05 (5%) and standard deviation 0.04. The correlation between R_s and R_b is 0.24. If you place a fraction w of your money in the stock fund and the rest, $1 - w$, in the bond fund, then the return on your investment is $R = wR_s + (1 - w)R_b$.

Suppose that $w = 0.47$. Compute the mean and standard deviation of R.

The mean is ______. (*Round your response to three decimal places.*)

The standard deviation is ______. (*Round your response to three decimal places.*)

Suppose that $w = 0.71$. Compute the mean and standard deviation of R.

The mean is ______. (*Round your response to three decimal places.*)

The standard deviation is ______. (*Round your response to three decimal places.*)

What value of w makes the mean of R as large as possible?

$w =$ ______ maximizes μ. (*Round your response to two decimal places.*)

What is the standard deviation of R for this value of w?

$\sigma =$ ______ for this value of w. (*Round your response to two decimal places.*)

What is the value of w that minimizes the standard deviation of R?

$w =$ ______ minimizes the standard deviation of R. (*Round your response to two decimal places.*)

Answers 0.064

0.044

0.071

0.055

1

0.07

0.18

ID: Exercise 2.22

---

24. The correlation between X and Y:

- **A.** is given by $\text{corr}(X,Y) = \dfrac{\text{cov}(X,Y)}{\text{var}(X)\text{var}(Y)}$.
- **B.** is the covariance squared.
- **C.** can be calculated by dividing the covariance between X and Y by the product of the two standard deviations.
- **D.** cannot be negative since variances are always positive.

Answer: C. can be calculated by dividing the covariance between X and Y by the product of the two standard deviations.

ID: Test A Ex 2.3.5

---

25. The following table gives the joint probability distribution of the speed of a car (C) and the incidence of an accident (A), where $A = 0$ when no accidents have occured and $A = 1$ when an accident has occured:

	C = 30 kmph	**C = 45 kmph**	**C = 90 kmph**
A = 0	0.30	0.07	0.10
A = 1	0.05	0.13	0.35

Complete the following table with the marginal distributions of the speed of a car (C) and the incidence of an accident (A).

	C = 30 kmph	**C = 45 kmph**	**C = 90 kmph**	**Marginal probability (A)**
A = 0	0.30	0.07	0.10	
A = 1	0.05	0.13	0.35	
Marginal probability (C)				

Complete the following table with the conditional distribution of the speed of a car given the incidence of an accident, i.e., C given A.

	C = 30 kmph	**C = 45 kmph**	**C = 90 kmph**	**Total**
$\Pr(C\|A = 0)$				
$\Pr(C\|A = 1)$				

(*Round your answers to two decimal places.*)

The conditional expectation of the speed of the car, given no accidents take place, $E(C|A = 0)$, is ______ kmph.

And, the conditional expectation of the speed of the car, given an accident takes place, $E(C|A = 1)$, is ______ kmph.

(*Round your answers to two decimal places.*)

The probability $\Pr(C = 90|A = 1) =$ ______ and the probability $\Pr(C = 90) =$ ______.

The probability of having an accident (A) and the speed of a car (C) are (1) ______ events because (2) ______.

(1) ○ dependent
○ independent

(2) ○ $\Pr(C = 90|A = 1) = \Pr(C = 90)$
○ $\Pr(C = 90|A = 1) \neq \Pr(C = 90)$

Answers 0.47

0.53

0.35

0.20

0.45

1.00

0.64

0.15

0.21

1.00

0.09

0.25

0.66

1.00

44.85

73.35

0.66

0.45

(1) dependent

(2) $\Pr(C = 90 | A = 1) \neq \Pr(C = 90)$

ID: Concept Exercise 2.3.1

26. Let there be two players in a game, Player 1 and Player 2. Consider a jar containing 5 snakes. 3 of the snakes in the jar are venomous, while the remaining 2 are non-venomous. In the game, both the players have to put their hand in the jar one after the other and pick a snake out. Each snake, if picked out of the jar, will bite the player's hand. The event of picking a venomous snake, or equivalently, a venomous snake's bite will earn the player zero points. On the other hand, the event of picking a non-venomous snake, or equivalently, a non-venomous snake's bite will earn the player one point.

Let X denote Player 1's pick and let Y denote Player 2's pick. Suppose Player 1 is the first to pick out a snake.

The expected value of Player 1's pick is: $E(X) =$ ______.

(*Express your answer as a fraction or round your answer to two decimal places.*)

The expected value of Player 2's pick is: $E(Y) =$ ______.

(*Express your answer as a fraction or round your answer to two decimal places.*)

Which of the following statements describes the relationship between $E(X)$ and $E(Y)$ in this example?

- **A.** $E(X)$ and $E(Y)$ are equal, so the order in which the players pick a snake is irrelevant.
- **B.** $E(X)$ is greater than $E(Y)$ because Player 1 has an advantage of picking first.
- **C.** $E(Y)$ is greater than $E(X)$ as there is a greater possibility that Player 1 picks up a venomous snake.
- **D.** $E(X)$ and $E(Y)$ are independent of each other. Their values do not reflect anything about their relationship.

Answers $\frac{2}{5}$

$\frac{2}{5}$

A. $E(X)$ and $E(Y)$ are equal, so the order in which the players pick a snake is irrelevant.

ID: Concept Exercise 2.3.2

27. The table gives the joint probability distribution of the number of sports an individual plays (X) and the number of times she may get injured while playing (Y).

	X = 1	**X = 2**	**X = 3**
Y = 4	0.12	0.08	0.15
Y = 3	0.05	0.05	0.06
Y = 2	0.10	0.03	0.15
Y = 1	0.15	0.04	0.02

The covariance between X and Y, σ_{XY}, is [ ].

(*Round your answer to two decimal places. Enter a minus sign if your answer is negative.*)

The correlation between X and Y, corr(X, Y), is [ ].

(*Round your answer to two decimal places. Enter a minus sign if your answer is negative.*)

An increase in the number of sports an individual plays will tend to (1) ________ the number of times she may get injured while playing.

(1) ○ increase
○ decrease

Answers 0.23

0.22

(1) increase

ID: Concept Exercise 2.3.3

---

28. Suppose that $Y_1, \ldots, Y_n$ are i.i.d. random variables with a $N(\mu_Y, \sigma_Y^2)$ distribution. How would the probability density of $\bar{Y}$ change as the sample size n increases?

Hint: Think about the law of large numbers.

○ **A.** As the sample size increases, the variance of $\bar{Y}$ increases. So, the distribution of $\bar{Y}$ becomes highly concentrated around μ_Y.

○ **B.** As the sample size increases, the variance of $\bar{Y}$ decreases. So, the distribution of $\bar{Y}$ becomes less concentrated around μ_Y.

○ **C.** As the sample size increases, the variance of $\bar{Y}$ increases. So, the distribution of $\bar{Y}$ becomes less concentrated around μ_Y.

○ **D.** As the sample size increases, the variance of $\bar{Y}$ decreases. So, the distribution of $\bar{Y}$ becomes highly concentrated around μ_Y.

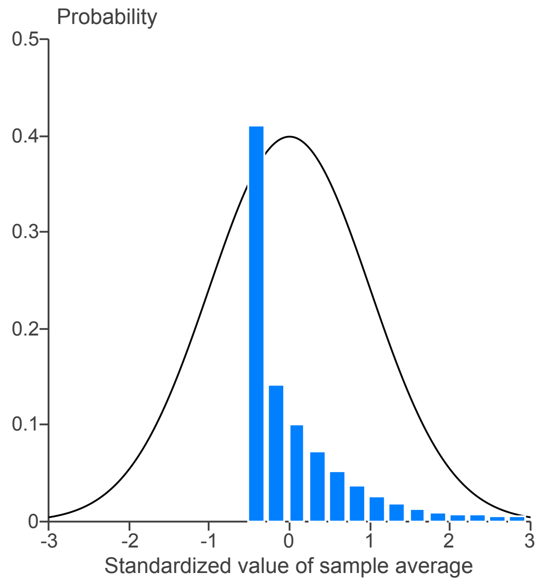
Answer: D.

As the sample size increases, the variance of $\bar{Y}$ decreases. So, the distribution of $\bar{Y}$ becomes highly concentrated around μ_Y.

ID: Review Concept 2.5

---

29. Suppose that Y_1, ..., Y_n are i.i.d. random variables with the probability distribution given in the figure below.



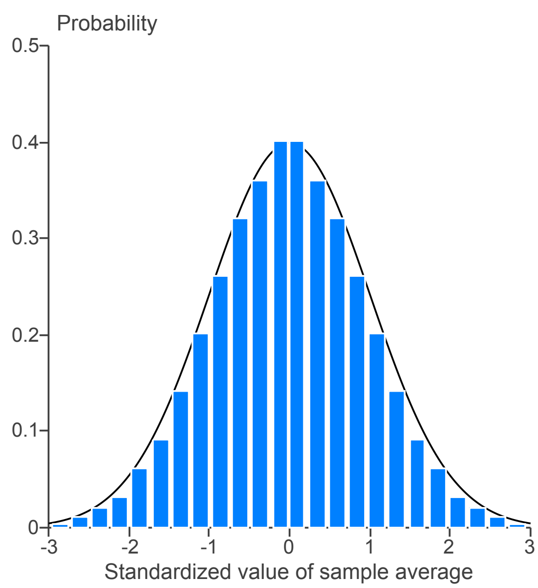
n = 25

Suppose further that you want to calculate $\Pr\left(\bar{Y} \le 0.1\right)$.

Would it be reasonable to use the normal approximation if n = 25?

- **A.** Yes.
- **B.** No.

Now suppose that with n = 1000 the distribution looks as follows.



n = 1000

Would it be reasonable to use the normal approximation now?

- **A.** Yes.
- **B.** No.

Answers B. No.

A. Yes.

ID: Review Concept 2.6

---

30. Suppose Y is a random variable with $\mu_Y = 0$, and $\sigma_Y^2 = 1$, skewness = 0, and kurtosis = 100. Explain why n random variables drawn from this distribution might have some large outliers.

- **A.** There might be some outliers because the skewness of the distribution equals 0.
- **B.** There might be some outliers because the kurtosis of the distribution equals 100.
- **C.** There might be some large outliers because both, the population variance σ_Y^2 and standard deviation σ_Y, are equal to 1.
- **D.** There might be some large outliers because the population μ_Y equals 0.

Answer: B. There might be some outliers because the kurtosis of the distribution equals 100.

ID: Review Concept 2.7

---

31. Compute the following probabilities:

If Y is distributed $N(-8,9)$, $\Pr(Y \leq -7) =$ ______. (*Round your response to four decimal places.*)

If Y is distributed $N(3,9)$, $\Pr(Y > 3) =$ ______. (*Round your response to four decimal places.*)

If Y is distributed $N(7,9)$, $\Pr(4 \leq Y \leq 12) =$ ______. (*Round your response to four decimal places.*)

Answers 0.6293

0.5000

0.7939

ID: Exercise 2.10

---

32. The following table contains data on the joint distribution of age (*Age*) and average hourly earnings (*AHE*) for 25 to 34 year-old full-time workers with an educational level that exceeds a high school diploma in 2012. Download the data from the table by clicking the *download table* icon . A detailed description of the variables used in the dataset is available here . Use a statistical package of your choice to answer the following questions.

Compute the marginal distribution of *Age*.

Marginal distribution of *Age*

	***Age* (years)**								
	25	**26**	**27**	**28**	**29**	**30**	**31**	**32**	
=									

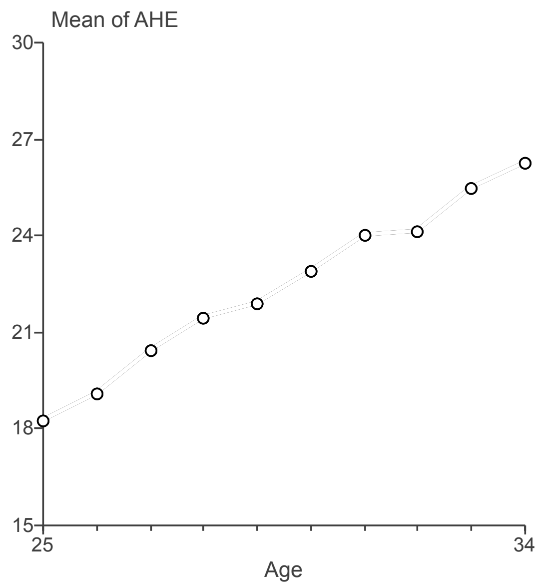
(*Round your response to four decimal places*)

Compute the mean of *AHE* for *Age* = 30; that is, compute, $E(AHE \mid Age = 30)$.

$$E(AHE \mid Age = 30) = \square$$

(*Round your response to four decimal places*)

Below is the plot of the mean of *AHE* versus *Age*.



Are average hourly earnings and age related?

- **A.** No, the mean of *AHE* and *Age* do not seem to be related.
- **B.** Yes, the mean of *AHE* and *Age* appear to be negatively related.
- **C.** Yes, the mean of *AHE* and *Age* appear to be positively related.

Use the law of iterated expectations to compute the mean of *AHE*; that is compute $E(AHE)$.

$$E(AHE) = \square$$

(*Round your response to four decimal places*)

Compute the variance of *AHE*; that is compute var(*AHE*).

$$\mathrm{var}(AHE) = \square$$

(*Round your response to four decimal places*)

Compute the covariance between *AHE* and *Age*; that is compute cov(*AHE*, *Age*).

$$\mathrm{cov}(AHE, Age) = \square$$

(*Round your response to four decimal places*)

Compute the correlation between *AHE* and *Age*; that is compute corr(*AHE*, *Age*).

corr(*AHE*, *Age*) = ☐

(*Round your response to four decimal places*)

Answers 0.0754

0.0797

0.0947

0.0953

0.0983

0.1105

0.1109

0.1134

0.1134

0.1085

22.8821

C. Yes, the mean of *AHE* and *Age* appear to be positively related.

22.6914

154.1000

6.7205

0.1932

ID: Empirical Exercise 2.1

---

33. Assume that *Y* is normally distributed $N(\mu, \sigma^2)$. Moving from the mean (μ) 1.96 standard deviations to the left and 1.96 standard deviations to the right, then the area under the normal p.d.f. is:

- **A.** 0.95.
- **B.** 0.33.
- **C.** 0.05.
- **D.** 0.67.

Answer: A. 0.95.

ID: Test B Ex 2.4.5

---

34. X_1, X_2, X_3, and X_4 are normally distributed random variables: $X_1 \sim N(0,0)$, $X_2 \sim N(0,1)$, $X_3 \sim N(1,0)$, and $X_4 \sim N(1,1)$.

The variable(s) which follow(s) a standard normal distribution is (are) (1) ________.

Y_1, Y_2, and Y_3 are normally distributed random variables. Use the normal cumulative distribution function to answer the following questions.

If $Y_1 \sim N(6,1)$, $\Pr(Y_1 \le 5.5) =$ ______.

If $Y_2 \sim N(2,25)$, $\Pr(Y_2 > 10) =$ ______.

If $Y_3 \sim N(29,16)$, $\Pr(20 \le Y_3 \le 40) =$ ______.

(*Round your answers to four decimal places.*)

Out of the above, the variable(s) which has (have) a skewness value of zero and a kurtosis value of 3 is (are) (2) ________.

(1)
- ○ X_1
- ○ X_2
- ○ X_3
- ○ X_4
- ○ all of the above
- ○ none of the above

(2)
- ○ X_1
- ○ X_2
- ○ X_3
- ○ X_4
- ○ all of the above
- ○ none of the above

Answers (1) X_2

0.3085

0.0548

0.9848

(2) all of the above

ID: Concept Exercise 2.4.1

35. In a given population for beverage drinkers, an individual's per kg expenditure on tea (*T*) and their per kg expenditure on coffee (*C*) have a bivariate normal distribution with covariance 0.15. An individual's per kg expenditure on tea is distributed with mean $2.85 and variance 0.16. An individual's per kg expenditure on coffee is distributed with mean $2.32 and variance 0.09.

If each individual in the population drinks 3 kg of tea and 2 kg of coffee, the mean total expenditure on beverages is $[ ] with a variance of [ ].

If *T* and *C* have a bivariate normal distribution with covariance zero, the mean total expenditure on beverages is $[ ] with a variance of [ ].

If *X* and *Y* have a bivariate distribution with covariance zero, this implies that the variables show (1) ________.

(1)
- ○ consistency
- ○ skewness
- ○ independence

Answers 13.19

3.60

13.19

1.80

(1) independence

ID: Concept Exercise 2.4.2

36. Let A, B, C, D, E, and F be independent standard normal random variables. Identify the distributions that will be followed by the variables P, Q, and R.

Variable	Distribution
$P = A^2 + B^2 + C^2$	(1) ______
$Q = \dfrac{F}{\sqrt{\left(D^2 + E^2\right)/2}}$	(2) ______
$R = \dfrac{\left(A^2 + B^2 + C^2\right)/3}{\left(D^2 + E^2\right)/2}$	(3) ______

Compute the following probabilities.

threetwo

Group 1: (*Round your answers to three decimal places.*)

If X is distributed χ^2_{30}, $\Pr(X > 50.89) =$ ______.

If X is distributed t_{30}, $\Pr(X > 2.04) =$ ______.

Group 2: (*Round your answers to two decimal places.*)

If X is distributed χ^2_{15}, $\Pr(X \le 25) =$ ______.

If X is distributed $F_{15,\infty}$, $\Pr(X \le 1.67) =$ ______.

Group 3: (*Round your answers to three decimal places.*)

If X is distributed $F_{7,\infty}$, $\Pr(X \le 2.01) =$ ______.

If X is distributed t_7, $\Pr(X \le 2.36) =$ ______.

Group 4: (*Round your answers to two decimal places.*)

If X is distributed $F_{10,2}$, $\Pr(X \le 19.40) =$ ______.

If X is distributed $F_{2,10}$, $\Pr(X \le 2.92) =$ ______.

Group 5: (*Round your answers to three decimal places.*)

If X is distributed $N(0,1)$, $\Pr(X > 1.86) =$ ______.

If X is distributed t_8, $\Pr(X > 1.86) =$ ______.

The probabilities of (4) ______ are equal because (5) ______.

(1)
- F distribution
- Normal distribution
- Chi − squared distribution
- t distribution

(2)
- t distribution
- Normal distribution
- F distribution
- Chi − squared distribution

(3)
- t distribution
- Normal distribution
- F distribution
- Chi − squared distribution

(4)
- Group 2
- Group 3
- Group 4
- Group 1
- Group 5

(5)
- the F − distribution is symmetric
- $F_{m,\infty}$ distribution is the distribution of χ^2_m / m
- the distribution of χ^2_m is the same as the distribution of t_m for large samples
- if $X \sim t_n$, then $X \sim F_{1,n}$
- the t − distribution is always approximately equal to the standard normal distribution

Answers (1) Chi − squared distribution

(2) t distribution

(3) F distribution

0.01

0.025

0.95

0.95

0.95

0.975

0.95

0.90

0.031

0.05

(4) Group 2

(5) $F_{m,\infty}$ distribution is the distribution of χ^2_m / m

ID: Concept Exercise 2.4.3

---

37. In any year, the weather can inflict storm damage to a home. From year to year, the damage is random. Let Y denote the dollar value of damage in any given year. Suppose that in 95% of the years $Y = \$0$, but in 5% of the years $Y = \$20{,}318$.

The mean of the damage in any year is \$_____. (*Round your response to two decimal places.*)

The standard deviation of the damage in any year is \$_____. (*Round your response to two decimal places.*)

Consider an "insurance pool" of 100 people whose homes are sufficiently dispersed so that, in any year, the damage to different homes can be viewed as independently distributed random variables. Let $\overline{Y}$ denote the average damage to these 100 homes in a year.

$E(\overline{Y})$, the expected value of the average damage $\overline{Y}$, is \$_____. (*Round your response to two decimal places.*)

The probability that $\overline{Y}$ exceeds \$2,000 is _____. (*Round your response to four decimal places.*)

Answers 1,015.9

4,428.21

1,015.9

0.0132

ID: Exercise 2.18

---

38. In any year, a person can suffer from a minor fracture. From year to year, the number of people seeking treatment for such fractures is random. Let Y denote the treatment expenditure for a minor fracture in any given year. Suppose that in 92% of the years $Y = \$0$, but in 8% of the years $Y = \$4,000$.

The mean treatment expenditure for a minor fracture in any year is $ [ ], and the standard deviation of the treatment expenditure for a minor fracture in any year is $ [ ].

(*Round your answers to two decimal places.*)

Consider a group of 484 people whose lives, homes, and occupations are sufficiently dispersed so that, in any year, the treatment expenditure for a minor fracture of different persons can be viewed as independently distributed random variables. Let $\bar{Y}$ denote the average treatment expenditure for a minor fracture of these 484 persons in a year.

The expected value of the average treatment expenditure for a minor fracture, $E(\bar{Y})$, in any year is $ [ ], and the standard deviation of the average treatment expenditure for a minor fracture in any year is $ [ ].

(*Round your answers to two decimal places.*)

The probability that for 484 people $\bar{Y}$ exceeds $500 is [ ].

(*Round your answer to four decimal places.*)

Answers 320

1,085.17

320

49.33

0.0001

ID: Concept Exercise 2.6.1

---

39. A software firm has 120 employees. The average time spent by these employees on social networking sites in a day during office hours is 55 minutes. Suppose a random sample of 10 employees is selected from the firm, and their average time spent on social networking sites is calculated.

Considering that the time spent by any two employees is independent of each other, which of the following statements are true? (*Check all that apply*.)

☐ **A.** In this case, the sample mean is likely to be consistent for the population mean.

☐ **B.** The mean time spent by these randomly selected 10 employees will approximately be 55 minutes as the sample size is large.

☐ **C.** The mean time spent by these randomly selected 10 employees may not be 55 minutes as the sample size is small.

☐ **D.** In this case, the sample mean is likely to be not consistent for the population mean.

Answer: C. The mean time spent by these randomly selected 10 employees may not be 55 minutes as the sample size is small., D. In this case, the sample mean is likely to be not consistent for the population mean.

ID: Concept Exercise 2.6.2

---

40. Which of the following statements best describes what the central limit theorem states?

- **A.** Under general conditions, when n is large, the distribution of $\bar{Y}$ is well approximated by a standard normal distribution even if Y_i are not themselves normally distributed.
- **B.** Under general conditions, when n is large, the distribution of $\bar{Y}$ is well approximated by a normal distribution even if Y_i are not themselves normally distributed.
- **C.** Under general conditions, the mean of Y is the weighted average of the conditional expectation of Y given X, weighted by the probability distribution of X.
- **D.** Under general conditions, when n is large, $\bar{Y}$ will be near μ_Y with very high probability.

Answer: B.

Under general conditions, when n is large, the distribution of $\bar{Y}$ is well approximated by a normal distribution even if Y_i are not themselves normally distributed.

ID: Concept Exercise 2.6.3