Started on	Monday, 25 April 2022, 10:57 AM
State	Finished
Completed on	Wednesday, 27 April 2022, 10:39 PM
Time taken	2 days 11 hours

Question **1**

Correct

Marked out of 1.00

A survey was conducted to ask a small random sample of individuals 18 years and older attending a big soccer game whether they had "received unemployment insurance in the last 5 years" (the event of interest). The data are given in the table below. The codebook for the data says that 1 means yes, and 0 means no. The objective is to use these data to estimate the proportion of soccer fans in the larger population of soccer fans for which the event happened. After looking at the data, match the following questions.

ID	unemp	ID	unemp
1	1	10	0
2	0	11	0
3	0	12	1
4	0	13	1
5	1	14	0
6	0	15	0
7	1	16	0
8	0	17	0
9	0	18	1

The estimate of the probability that a randomly chosen adult among ALL adult soccer fans received unemployment insurance in the last 5 years is — 0.3334 ✔

Let Y be the number of soccer fans in a random sample of 18 soccer fans that received unemployment insurance in the past 5 years. Estimate the probability that 2 or less in the sample received the insurance. — 0.0328 ✔

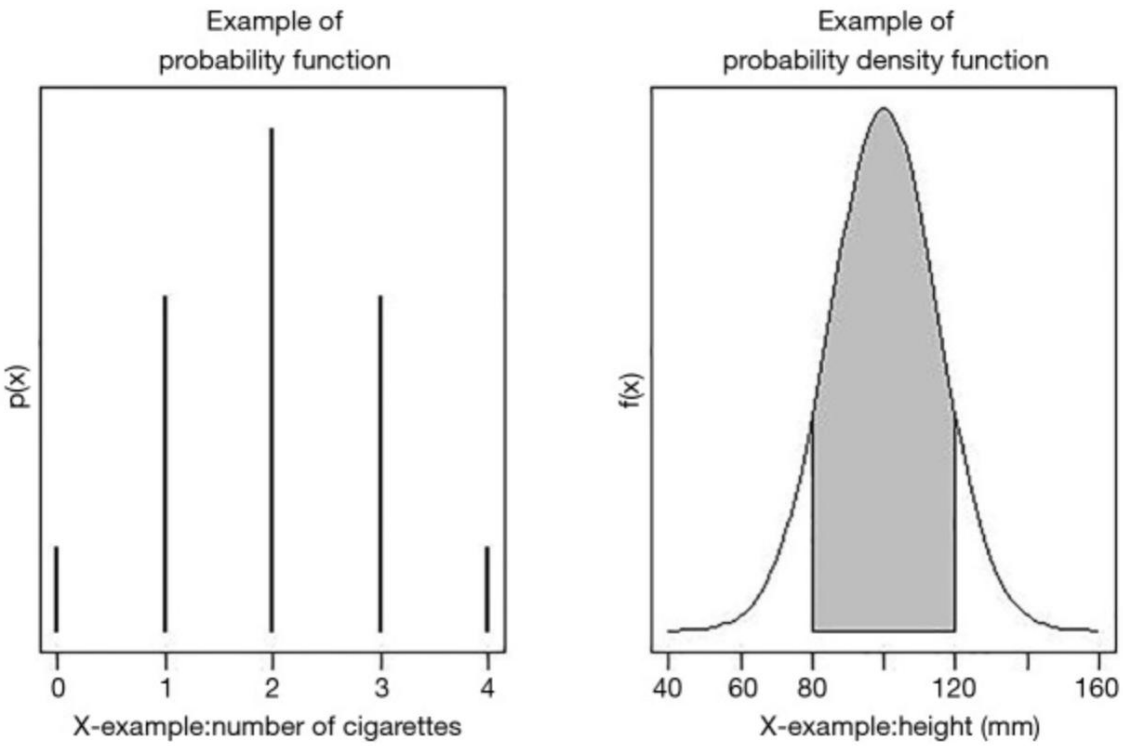
We estimate that in a random sample of 18 soccer fans we should expect this many that received unemployment insurance in the last year — 5.9999 ✔

The expected value is never what we actually get. A typical deviation is — square root of 3.99999 ✔

Question **2**

Partially correct

Marked out of 1.00



Pictured above are the plots of the pmf of a discrete random variable and the plot of the pdf of a continuous random variable. For each of the following variables, determine what might be an appropriate model to fit to data if you had data on it.

The number of electric scooters left per block of a Los Angeles street.	Poisson ✔
A metaanalysis of sequential clinical trials was done to obtain a data set containing a variable that measures, for each trial, how many patients that got treatment it took to get the first that got seriously ill with the treatment. Note: a sequential clinical trial is one in which the patients in the treatment group are not all given the treatment at once but rather sequentially: first one patient, observe, then another patient, observe, etc... A metaanalysis is an analysis of many research articles to gather information on all of them and conduct analysis of the results of all.	Geometric ✔
A variable measuring the number of students in a random sample of 20 from California who know what the "Bill of Rights" is	Binomial ✔
Body temperature of humans	Normal ✔
The duration of the period when an electric scooter is idle and awaiting the next customer	Exponential ✔
The proportion of hours in a given day that a train arrives late to a train station	Poisson ✖

Question 3

Correct

Marked out of 1.00

The Diabetes Control and Complications trial was a trial that had a lot of consequences for diabetes management. It established that intensive glucose measuring and constant management of insulin delivery helped decrease the long term effects of diabetes. A few lines and variables are given below. Match the variable with the parameter and model you would estimate using that data, meaning that you are assuming that the probability model would be the right one for that variable. Notice that we are using the notation used in the lectures and our textbook. For the "adult" variable I (meaning you) would estimate the parameter [p] ✔ of the [Bernoullli] ✔ model. With that parameter I then estimate a model for the number of patients that are adult in a random sample as large as that participating in the trial, which is a [Binomial] ✔ model. For "duration" which, according to the codebook, means duration of diabetes in months, I would estimate the parameter [mu] ✔ because this is a [Continuous] ✔ random variable that could be modeled as a [Gaussian] ✔ , since many factors affect the onset of diabetes. Perhaps I would also model age as [Gaussian] ✔ but I am not so sure of that, because it is possible that women and men (which are the only gender identified in this trial) have different distributions and it could be that the distribution of age is [bimodal] ✔

	group	phase	retbase	adult	age	duration
2	EXPERIMENT	2	SCND	0	17	178
3	STANDARD	2	SCND	1	29	142
4	STANDARD	2	SCND	1	35	175
5	EXPERIMENT	2	PRIM	0	14	31
6	EXPERIMENT	2	SCND	1	32	72
7	STANDARD	2	SCND	1	26	106
8	STANDARD	2	SCND	1	26	168
9	EXPERIMENT	2	SCND	1	28	147
10	EXPERIMENT	2	SCND	1	37	14
11	STANDARD	2	SCND	1	23	80
12	STANDARD	2	SCND	0	13	148
13	STANDARD	2	PRIM	0	13	30
14	STANDARD	2	SCND	1	21	126
15	STANDARD	2	SCND	1	27	116
16	EXPERIMENT	2	SCND	1	38	133
17	STANDARD	2	PRIM	1	37	38
18	EXPERIMENT	2	SCND	1	27	40
19	STANDARD	2	SCND	1	23	61
20	EXPERIMENT	2	SCND	0	17	77
21	EXPERIMENT	2	SCND	1	22	35
22	STANDARD	2	SCND	1	25	168
23	EXPERIMENT	2	SCND	0	14	71
24	EXPERIMENT	2	PRIM	1	28	27
25	STANDARD	2	SCND	0	15	107

[Gaussian] [Gaussian]

[geometric] [negative binomial] [Poisson] [Exponential]

Question **4**

Correct

Marked out of 1.00

A bakery has 5 ovens. At least 4 ovens must be working in order to meet customer demand on a given day. The probability of a particular oven working is 0.9. We want to find out the probability of meeting customer demand.

Select one:

- a. 0.9185 ✔
- b. 0.32805
- c. 0.59049
- d. 0.5781
- e. 0.7891

Question **5**

Correct

Marked out of 1.00

When circuit boards used in the manufacture of DVD players are tested, the long-run percentage of defectives is 5%. Let X = the number of defective boards in a random sample of size n = 25, so X ~ Bin(n=25; p=05).

Match the following:

$P(X \leq 2)$	0.873 ✔
$P(X \geq 5)$	0.007 ✔
$P(1 \leq X \leq 4)$	0.716 ✔
μ	1.25 ✔
σ	1.09 ✔
Probability that none of the 25 boards is defective	0.277 ✔

Question **6**

Correct

Marked out of 1.00

A recruiting firm finds that 20% of the applicants for a particular sales position are fluent in both English and Spanish. Applicants are selected at random from the pool and interviewed sequentially.

Suppose that the first applicant who is fluent in both English and Spanish is offered the position, and the applicant accepts. Suppose each interview costs $125. The expected value and standard deviation of the cost of interviewing until the job is filled. are, respectively,

Select one:

- a. $1000.45 and $125.18
- b. $625 and $559.017 ✔
- c. $500 and $100.921
- d. $250 and $131.1
- e. $2034.15 and $20.

Question **7**

Correct

Marked out of 1.00

The Center for Disease Control says that about 30% of high school students smoke tobacco (down from a high of 38% in 1997). Suppose you randomly select high school students to survey them on their attitude towards scenes of smoking in the movies. What is the probability that it takes less than three students surveyed to find the first smoker?

Select one:

- a. 0.657
- b. 0.51 ✔
- c. 0.3
- d. 0.7
- e. 0.77

Question **8**

Correct

Marked out of 1.00

A recruiting firm finds that 30 percent of the applicants for a certain industrial job have received advanced training in computer programming. Applicants are interviewed sequentially and selected at random from the pool. Suppose that the first applicant with advanced training is offered the position, and the applicant accepts. If each interview costs $30, the expected value and variance of the total cost of interviewing until the job is filled are given, respectively, by

Select one:

- a. 100 and 7000 ✔
- b. 3.33 and 7.777
- c. 50 and 9
- d. 200 and 1000
- e. 150 and 500

Question **9**

Correct

Marked out of 1.00

Let X be the number of multiple choice questions a student gets right on a 40-question test, when each question has 4 choices (and only one of the 4 choices is correct) and the student is completely guessing.

The random variable X is

Select one:

- a. A binomial random variable ✔
- b. A Poisson random variable
- c. A Bernoulli random variable
- d. A geometric random variable

Question **10**

Correct

Marked out of 1.00

A large stockpile of used pumps contains 20% that are currently unusable and need to be repaired. A repairman is sent to the stockpile with three repair kits. He selects pumps at random and tests them one at a time. If a pump works, he goes on to the next one. If a pump doesn't work, he uses one of his repair kits on it. Suppose that it takes 10 minutes to test whether a pump works, and 20 minutes to repair a pump that does not work. The expected value and variance of the total time it takes the repairman to use up his three kits are, respectively,

Select one:

- a. 210 and 6000 ✔
- b. 210 and 600
- c. 60 and 210
- d. 60 and 2000
- e. 100 and 1500

Question **11**

Correct

Marked out of 1.00

Consider a sequence of random variables

$$Y_1, Y_2, \ldots.$$

where each

$$Y_i$$

is Bernoulli. Random variable X equals the value of i such that

$$Y_i$$

is the first Y with value 1. The random variable X is

Select one:

- a. An exponential Bernoulli random variable
- b. A geometric random variable ✔
- c. A binomial random variable
- d. A Poisson random variable.

Question **12**

Correct

Marked out of 1.00

$$\sum_{y=0}^{n-1} y\binom{n-1}{y}p^y(1-p)^{n-1-y} + \sum_{y=0}^{n-1}\binom{n-1}{y}p^y(1-p)^{n-1-y}$$

is equal to

Select one:

- a. 5
- b. (n-1)p +1 ✔
- c. 1
- d. 0
- e. (n+1)

Question **13**

Correct

Marked out of 1.00

Which of the following statements is correct?

Select one or more:

- [ ] a. The sample space corresponding to a binomial random variable always has simple outcomes that consist of a sequence of three Bernoulli trials, e.g., ddd, ddu, .. that is, n is always 3 and p is anything.
- [x] b. The difference between a binomial experiment and a geometric experiment where we screen individuals until we find the first person with lung cancer is that in a binomial experiment, all outcomes consist of sequences with a fixed number of Bernoulli trials, whereas in the lung experiment, the number of Bernoulli trials is different for each outcome ✔
- [x] c. The Expected value of a binomial random variable is n times the expected value of a Bernoulli random variable. ✔

Question **14**

Correct

Marked out of 1.00

In the article posted in Module 4, authored by Sommerfeld (with title "The Binomial and Hypergeometric Probability Distributions in Jury Selection," the authors show how the Binomial and the Hypergeometric models can be used to show that a jury panel is biased.

A way to do that is to use the population parameter p (probability that an individual in the population has the characteristic of interest) then calculate the probability that in a random sample of size n, we would find a given number of individuals with that characteristic.

For example, if in San Joaquin County 5% of the population are African American, and we choose a random sample of size 105 from that population, using the Binomial model we can calculate that the Probability of having 0 African American by chance in the sample is 0.00458, using the binomial formula. You can use the app posted in Module 5 to calculate. https://homepage.divms.uiowa.edu/~mbognar/applets/bin.html

The fact that by chance you are not likely to see 0 African Americans in the sample indicates that the panel is biased, not random (of course, statisticians could come up with additional tests to be sure, but probability that low says that if the random sample is really random, that is not possible).

With that information given above, what would be the expected number of African Americans in that sample of 105 from San Joaquin County?

- [x] a. 5.25 ✔
- [ ] b. 0
- [ ] c. 1
- [ ] d. 1

Question **15**

Correct

Marked out of 1.00

CHAPTER 5 TEXTBOOK Section 5.18- Exercise 20

Let X be the number of bacterial colonies per cubic centimeter, a Poisson random variable with expected value 3.
(i) What is the probability that there is at least one bacterial colony in a randomly chosen cubic centimeter?
(ii) What is the probability that in five randomly chosen cubic centimeters there is at least one cubic centimeter where there is at least one bacterial colony?
(iii) How many cubic centimeters must be observed for the probability of observing at least one with at least one bacterial colony to be 0.95?

What is the probability that there is at least one bacterial colony in a randomly chosen cubic centimeter? 0.9502 ✔

What is the probability that in five randomly chosen cubic centimeters there is at least one cubic centimeter where there is at least one bacterial colony? 0.9999997 ✔

How many cubic centimeters must be observed for the probability of observing at least one with at least one bacterial colony to be 0.95? 1 ✔

Question **16**

Correct

Marked out of 1.00

The number alpha particles emitted by a radioactive substance has expected value of 12 per square centimeter. If two 1-square centimeter samples are independently selected, find the probability that two received 4 alpha particles. How many 1-squarecentimeter samples should be selected to establish a probability of approximately 0.95 that at least one will contain one or more alpha particles?

If two 1-square centimeter samples are independently selected, find the probability that two received 4 alpha particles. approximately 0 ✔

How many 1-squarecentimeter samples should be selected to establish a probability of approximately 0.95 that at least one will contain one or more alpha particles? approximately 1 ✔

Question **17**

Correct

Marked out of 1.00

CHAPTER 5-Section 5.18-Exercise 18

Assume that 13% of people are left-handed. If we select five people at random, find the probability of each outcome below:

a. The first lefty is the fifth person chosen

b. There are some lefties among the five people

c. The first lefty is the second or third person

d. There are exactly three lefties in the group

e. There are no more than three lefties in the group

The first lefty is the fifth person chosen	0.07448 ✔
There are some lefties among the five people	0.5016 ✔
The first lefty is the second or third person	0.2115 ✔
There are no more than three lefties in the group	0.9987 ✔
There are exactly 3 lefties	0.01663 ✔

Question **18**

Correct

Marked out of 1.00

A resident of Boston spends the Summers in the Grand Tetons, Wyoming. Every day there is expectation that a moose may pass in front of the house. Moose are wild animals that live around that area. The daily sighting (number of times seen) of moose has the following probability mass function, where X is the daily number of sightings:

x	0	1	2
P(X=x)	0.1	0.5	0.4

If this Boston resident spends 5 days in the Grand Tetons, how many moose are expected to be seen in total in the 5 days? There is uncertainty, of course, the true number seen in total could be different from the expected value. By how much will the total number of sightings depart from the expected value, on average?

Select one:

a. 6.5 and $\sqrt{2.04992}$, respectively ✔

b. 5 and 1 respectively

c. 3.71293 and $\sqrt{0.1076261}$, respectively

d. 4.01411 and $\sqrt{2.31121}$ respectively

Question **19**

Correct

Marked out of 1.00

A resident of Boston spends the Summers in the Grand Tetons, Wyoming. Every day there is expectation that a moose may pass in front of the house. Moose are wild animals that live around that area. The daily sighting (number of times seen) of moose has the following probability mass function, where X is the daily number of sightings:

x	0	1	2
P(X=x)	0.1	0.5	0.4

If this Boston resident spends 5 days in the Grand Tetons, what is the probability that there will be at least one sighting every single day?

Select one:

- a. 0.59049 ✔
- b. 0.9
- c. 0.45
- d. 0.18

Question **20**

Correct

Marked out of 1.00

(Exercise 1, Sections 7.9.4) Wires manufactured for a certain computer system are specified to have a resistance of between 0.10 and 0.17 ohms. The actual measured resistances of the wires produced by company A have a normal probability density distribution, with expected value 0.13 ohms and standard deviation 0.005 ohms. If three independent such wires are used in a single system and all are selected from company A, what is the probability that they all will meet the specifications?

Select one:

- a. approx 0
- b. approx 1 ✔
- c. approx 0.5
- d. approx 0.315

Question **21**

Correct

Marked out of 1.00

(Exercise 11, in Sections 7.9.4) The weight of anodized reciprocating pistons produced by a company follows a Gaussian distribution with

$$\mu = 10$$

lb and standard deviation

$$0.2$$

lb. A sampling inspection scheme designed by the quality control engineers calls for rejecting the heaviest 2.5% of the pistons. What weight, in pounds, determines the overweight classification?

Select one:

- a. 9.995
- b. 10.392 ✔
- c. 10.2
- d. 10.6

Question **22**

Correct

Marked out of 1.00

(Exercises 5, in Sections 7.9.4) If

$$X$$

is a normal random variable with parameters

$$\mu = 3$$

and

$$\sigma^2 = 9$$

, find

$$P(|X - 3| > 6)$$

.

Select one:

- a. 0.9545
- b. 0.0003
- c. 0.001
- d. 0.0455 ✔

Question **23**

Correct

Marked out of 1.00

(Exercise 9, in Sections 7.9.4) Family branding occurs when a firm applies one brand name to its entire product line, such as Levi'€ s. Individual branding occurs when a firm uses individual brand names for its products, for example, Procter & Gambleâ€'s Pringles, Crisco, and Tide. GSP Inc. is trying family branding for a new toothpaste in 20 test cities. The mean and standard deviation of units sold per week are

$$2,250$$

and

$$250$$

respectively. GSP is also test marketing the toothpaste using individual branding in 20 similar cities. The mean and standard deviation in units sold per week are

$$2,250$$

and

$$500$$

. GSP will select the strategy that maximizes its chance of selling at least

$$2,350$$

units per week. This will ensure that it meets its return on the project'€ s investment goal. Which marketing approach-family or individual branding-should GSP select? Assume sales are normally distributed.

Select one:

- a. individual branding ✔
- b. family branding

Question **24**

Correct

Marked out of 1.00

Joe reads that one out of four eggs contains salmonella bacteria, so he never uses more than three eggs in cooking. If eggs do or do not contain salmonella independently of each other, the number of contaminated eggs when Joe uses three chosen at random has the distribution

Select one:

- a. binomial with n=4 and p=1/4
- b. binomial with n=3 and p=1/4 ✔
- c. binomial with n=3 and p = 1/3
- d. hypergeometric with total number of eggs equal 4 and a random sample of 2

Question **25**

Correct

Marked out of 1.00

Systolic blood pressure in normal healthy individuals is normally distributed with

$$\mu = 120$$

and

$$\sigma = 10$$

mm Hg. A systolic blood pressure of 136.45 mm Hg is at what percentile (approximately)?

Select one:

- a. 95th percentile ✔
- b. 25th percentile
- c. 88th percentile
- d. 69th percentile