

**BIOSTATS 540 - Introductory Biostatistics
Fall 2022**

Exam 2

**Unit 4 – Probabilities in Epidemiology
Unit 5 – Populations and Samples
Unit 6 – Bernoulli and Binomial Distributions
Unit 7 – Normal Distribution**

Due: Monday November 14, 2022

**Last Date for Submission with -10 points: Wednesday November 16, 2022
Last Date for Submission with Credit (-20 points): Monday November 21, 2022**

This is an “open book” “take-home” exam. You are welcome to use any reference materials you wish. You are welcome to use the computer as you wish, too. However, you **MUST** work this exam **by yourself** and you may **not** consult with anyone (except me!).

How to Submit Your Exam

Please upload a ***SINGLE PDF***
of your completed exam to the Blackboard Learn ASSIGNMENT tab
no later than **Monday November 14, 2022 (11:59 pm)**.

... unless you are opting to submit your exam late per the late submissions policy above.

**BIOSTATS 540 - Introductory Biostatistics
Fall 2022**

Exam 2

Unit 4 – Probabilities in Epidemiology

Unit 5 – Populations and Samples

Unit 6 – Bernoulli and Binomial Distributions

Unit 7 – Normal Distribution

Due: Monday November 14, 2022

Last Date for Submission with -10 points: Wednesday November 16, 2022

Last Date for Submission with Credit (-20 points): Monday November 21, 2022

Signature

This is to confirm that in completing this exam, I worked independently and did not consult with anyone.

Signature: _____

Printed Name: _____

Date: _____

10/31/2022

Dear BIOSTATS 540 Fall 2022

Like exam 1, the points on this exam total 105

Thus, you can lose 5 points in your work and still score 100 on this test.
Please note, however: The maximum score you can earn is 100.

1. (10 points total)

In a certain study, a sample of persons remanded into custody (incarcerated) in a certain jail were given a detailed interview. From this interview, researchers were able to ascertain who did and who did not take drugs at the time of incarceration. The following table shows the research findings for cannabis and it shows whether the prison’s screening procedure detected cannabis use. For this question, ***assume that the research findings are truth.***

Prison Screening Result	Research Finding	
	User	Non-user
User	49	6
Non-user	201	118

1a. (2 points)

Treating the research findings as truth, by hand, calculate the ***sensitivity*** of the prison screening procedure. Show your work.

1b. (2 points)

Treating the research findings as truth, by hand, calculate the ***specificity*** of the prison screening procedure. Show your work.

1c. (2 points)

Treating the research findings as truth, by hand, calculate the ***predictive value positive***. Show your work.

1d. (2 points)

Treating the research findings as truth, by hand, calculate the ***predictive value negative***. Show your work.

1e. (2 points)

Calculate the overall probability of concordance (**Note: Concordance refers to both procedures yielding the same result**). Show your work.

2. (10 points total)

The following table is from a case-control study into maternal smoking during pregnancy and Down syndrome. It shows the basic characteristics of mothers giving birth to babies with Down syndrome (cases), and without Down syndrome (controls).

		Cases (n=775)		Controls (n=7750)	
		Number	%	Number	%
Smoking during pregnancy, among					
Age < 35 years					
	Yes	112	20.0	1411	20.2
	No	421	75.0	5214	74.6
	Unknown	28	5.0	363	5.2
Age ≥ 35 years					
	Yes	15	7.0	108	14.2
	No	186	86.9	611	80.2
	Unknown	13	6.1	43	5.6

2a. (1 point)

Use the information in the table to construct a 2x2 table for women age < 35. For purposes of this question, exclude from analysis any woman for whom smoking status is Unknown.

2b. (1 point)

Use the information in the table to construct a 2x2 table for women age ≥ 35. Again, for purposes of this question, exclude from analysis any woman for whom smoking status is Unknown.

2c. (3 points)

Consider your 2x2 table for women age < 35. For these women, among mothers giving birth to babies with Down syndrome, by hand, calculate the odds that they had smoked during pregnancy. Again for these women, among mothers giving birth to a healthy baby, calculate by hand the odds that they had smoked during pregnancy. Finally, use your two odds calculations to obtain the odds ratio. In 1-2 sentences, explain the odds ratio in lay terms.

2d. (3 points)

Repeat the calculations you did to answer question 2c, this time for women age ≥ 35. In 1-2 sentences, explain the odds ratio in lay terms.

2e (2 points)

Finally, in 1-2 more sentences, what is your interpretation of the comparison of the two odds ratios? That is, what does this comparison suggest about age as a risk factor for delivering a Down syndrome baby?

3. (10 points total)**3a. (2 points)**

TRUE or FALSE: The number of sixes in 20 throws of a die follows a binomial distribution.

3b. (2 points)

TRUE or FALSE: The weight of a human follows a binomial distribution.

3c. (2 points)

TRUE or FALSE: The number of a random sample of patients who respond to a certain treatment follows a binomial distribution.

3d. (2 points)

TRUE or FALSE: The number of red cells in 1 ml of blood follows a binomial distribution.

3e. (2 points)

TRUE or FALSE: If a coin is tossed twice in succession, the number of tails follows a Binomial distribution.

4. (10 points total)

4a. (2 points)

A citrus farmer has observed the following probability distribution for the number of oranges per tree. How many oranges does he/she expect on average? (Hint: This question is asking you to calculate a statistical expectation)

Outcome, # oranges	25	30	35	40
Probability	.10	.40	.30	.20

4b. (2 points)

Consider again the probability distribution in question 4a. By hand, calculate the standard deviation of this distribution. Show your work.

4c. (2 points)

A manufacturer ships toasters in cartons of 20. In each carton, they estimate that each toaster has a 5% chance of being malfunctioning, thus requiring it to be sent back for repairs. What is the probability that in one carton, there will be exactly 3 toasters that need repair?

4d. (2 points)

A soccer team estimates that they will score on 8% of the corner kicks. In next week’s game, the team hopes to kick 15 corner kicks. What are the chances that they will score on exactly two of these opportunities?

4e. (2 points)

Consider again the setting in question 4d. If this soccer team has 200 corner kicks over the season, what are the chances that they will score more than 22 times?

5. (5 points)

Suppose that a fair coin is tossed several times, and you will win a dollar if more than 60% of the tosses land heads. Which is better for you: ten tosses or 100 tosses? Carry out the necessary calculations to obtain your answer and then explain it.

6. (20 points total)

Suppose it is known that the length of time (minutes) needed to service a car at a particular gas station is distributed Normal with $\mu = 4.5$ minutes and standard deviation $\sigma = 1.1$ minutes.

6a. (4 points)

What is the probability that a randomly selected car at this gas station will require more than 6 minutes of service or less than 5 minutes of service?

6b. (4 points)

What is the probability that a randomly selected car at this gas station will require between 3.5 and 5.6 minutes of service?

6c. (4 points)

What is the probability that a randomly selected car at this gas station will require at most 3.5 minutes of service?

6d. (4 points)

What is the servicing time such that only 5% of all cars at this gas station require more than this amount of time?

6e. (4 points)

What are the two servicing times, symmetric about the mean, within which 50% of all the cars at this gas station require this range of servicing time?

7. (5 points total)

Fill in the blank with the correct number needed to complete the sentence.

In a normal probability model frequency distribution, 33% of the outcomes fall between

$$\mu \pm \text{_____} \sigma$$

8. (10 points total)**8a. (5 points)**

Suppose it is known that the serum cholesterol levels of Wisconsin children are distributed normal with $\mu = 175$ mg/dL and standard deviation $\sigma = 30$ mg/dL. Suppose further that a “normal” value is defined as any value between two standard deviations of the mean. What are the values of these “normal limits”?

8b. (5 points)

Suppose it is known that IQ is distributed normal with $\mu = 100$ and standard deviation $\sigma = 15$. Now consider the average IQ of classes of 25 students. What are the population mean and variance of these class averages?

9. (20 points total)

A medium size bag of potato chips states on the bag that the weight of the potato chips is 10 ounces. A weight less than 10 ounces is considered “*underweight*”. Suppose that in producing medium size bags of potato chips, the actual weights are a population that is distributed normal with mean $\mu = 10.2$ ounces and standard deviation $\sigma = 0.12$ ounces.

9a. (5 points)

What fraction of all bags sold are underweight?

9b. (5 points)

What is the probability that the mean weight of a simple random sample of 3 bags is below the stated amount?

9c. (5 points)

What is the probability that, in a simple random sample of 3 bags, none are underweight?

9d. (5 points)

Suppose manufactured potato chips are bundled into 24-bag cases. You may assume each case represents a simple random sample of size 24. Consider the average weight of each 24-bag case of potato chips. Give two limits, symmetric about the mean, within which 99% of the mean weights lie.

10. (5 points)

The quality specification for a certain manufactured component is that the pressure at a certain point must not exceed 30 pounds. A manufacturer who would like to enter this market finds that he can make components with a mean pressure $\mu = 28$ lbs, but the pressure varies from one specimen to the next with a standard deviation of $\sigma = 1.6$ lbs.

What proportion of his/her specimens will fail to meet the quality specification?