Midterm Practice Problems

Ethan Levien

September 10, 2025

ableofcontents

0.1 Conditional and Marginal Probability

1. Given the joint probability table:

	$Y=0$	$Y=1$
$X=0$	0.1	0.2
$X=1$	0.3	0.4

   (a) Find $P(X=1)$. (b) Find $P(Y=0)$. (c) Find $P(X=1|Y=1)$. (d) Find $P(Y=1|X=0)$.

2. Let Z be a random variable with $P_Z(z)=0.2$ for $z=0$, 0.5 for $z=1$, 0.3 for $z=2$. Let $W|Z=z \sim$ Bernoulli$(z/4+1/4)$. Find $P(W=1|Z=2)$ and $P(W=1)$.

3. Suppose $A \sim$ Bernoulli(0.6) and $B|A \sim$ Bernoulli$(0.2A+0.3)$. Write the joint probability $P(A=a,B=b)$ for all $a,b \in \{0,1\}$.

4. Translate the model $X|Y \sim$ Bernoulli(p), $Y \sim$ Bernoulli(q) into a formula for $P(X=1)$ in terms of p and q.

5. Given the table:

	$B=0$	$B=1$
$A=0$	0.25	0.15
$A=1$	0.35	0.25

   Find $P(A=1)$, $P(B=1)$, $P(A=0|B=0)$, $P(B=1|A=1)$.

0.2 Python Syntax and Probability Models

1. Given $a = np.array([3,5,7,9])$, what is $a[2]$? What is $a[1:3]$?

2. Write Python code to generate 1000 samples from $X \sim$ Bernoulli(0.3).

3. Given the code:

```
for i in range(5):
print(i)
```

   What is the output?

4. Translate the following code into a probability model:

```
p = 0.7
samples = np.random.binomial(n=1, p=p, size=1000)
```

5. If $Y \sim$ Bernoulli(0.4) and $X|Y \sim$ Bernoulli($0.5Y + 0.2$), write Python code to generate 500 samples of (X, Y) pairs.

0.3 Statistical Inference and the Central Limit Theorem

1. If the sample mean of n iid samples from a distribution with variance σ^2 is used to estimate the population mean, what is the standard error of the sample mean?

2. How many samples are needed so that the standard error of the mean is less than 0.01 if $\sigma^2 = 1$?

3. Suppose you estimate the proportion of YES answers in a survey by sampling n people. If the true proportion is $p = 0.6$, what is the approximate standard error of your estimate for $n = 100$?

4. Explain the statement: "The sample mean converges to the population mean as n increases." What theorem justifies this?

5. If you want the probability that your sample mean differs from the true mean by more than 0.05 to be less than 5%, how many samples do you need (assume normality and $\sigma^2 = 1$)?