Statistics 2016 - Midterm

NAME: _____

You have 2 hours.
There are 7 questions, Each part of each question is worth 2 points.
You can use only a pen, a calculator, and a hand written cheat-sheet (both sides).

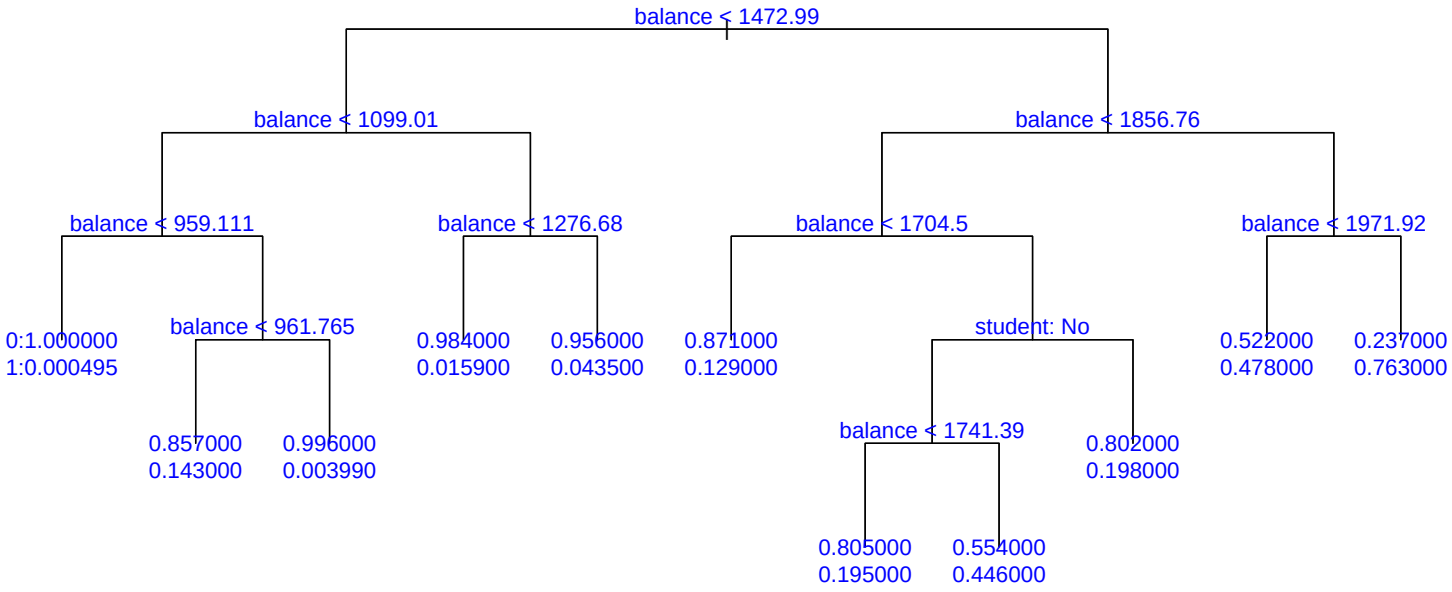
THE CONTENTS OF THIS EXAM ARE CONFIDENTIAL.
DO NOT DISCUSS THEM WITH ANYONE.

I pledge my honor that I have not violated the Honor Code
during this examination.

SIGNATURE: _____

1 Question

Here is the tree we had in the homework relating default to balance and student. Remember, the rule indicates when you go left.



1.1

What is the estimated probability of default for a student with a balance of 2000?

1.2

What is the estimated probability of default for a student with a balance of 1500?

2 Question

Let A be the random variable representing a return on an asset.

A is in "percent" so if A turns out to be 6, then the return is .06 and invested wealth would go up by a factor of 1.06 over the time period.

a	$P(A = a)$
0	.25
6	.5
12	.25

2.1

What is $P(A > 0)$, the probability of a positive return?

2.2

What is $E(A)$, the expected value of the return?

2.3

What is $Var(A)$, the variance of the return?

2.4

What is $sd(A)$, the standard deviation of the return?

3 Question

Suppose we think the probability the Leafs win their next game is .6.

Let $W \sim$ Bernoulli(.6) be the random variable which is 1 if Leafs win and 0 else.

Suppose I have a bet with a friend that pays me \$10 if the Leafs win and -\$10 if they lose. Let B be the outcome of the bet.

3.1

What is the linear function relating B to W?

3.2

What are the mean and standard deviation of W?

3.3

What are the mean and standard deviation of B?

3.4

What is the distribution of B? (give a little table).

4 Question

The table below gives the joint distribution of X and Y.

		X	
		0	1
Y	0	.24	.06
	1	.56	.14

4.1

What is $P(X = 1, Y = 0)$?

4.2

What is $P(Y = 1)$?

4.3

What is $P(Y = 1 \,|\, X = 1)$?

4.4

Are X and Y independent?

4.5

Is X a Bernoulli random variable?

4.6

What is $E(X)$?

4.7

Which is bigger, $Var(X)$ or $Var(Y)$?

4.8

Are X and Y IID?

5 Question

Suppose $R \sim N(10, 25)$ represents our beliefs about about the return on an asset over the next period.
R is in percent so that if R turns out to be 10, then wealth increases by a factor of 1.1.

5.1

What is $P(10 < R)$?

5.2

What is $P(0 < R < 20)$?

5.3

What is $P(5 < R < 15)$?

5.4

What is $P(0 < R)$, the probability of a positive return?

6 Problem

Suppose a person is randomly drawn from a large population and then tested for a disease.
Let $D = 1$ if the person has the disease and 0 otherwise.
Let $T = 1$ if the person tests positive and 0 otherwise.

Suppose
$P(D = 0) = .9$.
$P(T = 1 \mid D = 0) = .1$.
$P(T = 1 \mid D = 1) = .8$.

6.1

What is $P(D = 1)$?

6.2

What is $P(T = 1, D = 1)$?

6.3

What is $P(T = 1, D = 0)$?

6.4

What is $P(T = 1)$?

6.5

What is $P(D = 1 \mid T = 1)$?

6.6

Is this a very good test? Discuss.

7 Problem

Suppose we are sampling without replacement from a population of voters consisting of 6 Democrats and 4 Republicans.

Let X_i be 1 if the i^{th} sampled voter is Democrat and 0 otherwise.

Here is the joint distribution of (X_1, X_2, X_3).

x_1	x_2	x_3	$P(x_1, x_2, x_3)$
0	0	0	.033
0	0	1	.1
0	1	0	.1
0	1	1	.167
1	0	0	.1
1	0	1	.167
1	1	0	.167
1	1	1	.167

So, for example, the probability of getting dem, repub, repub, $= P(X_1 = 1, X_2 = 0, X_3 = 0) = .1$.

7.1

What should .033 + 3*.1 + 4*.167 be very close to?

7.2

What is $P(X_1 = 1)$?

7.3

What is $P(X_2 = 1)$?

7.4

What is $P(X_2 = 0, X_3 = 0)$?

7.5

What is $P(X_1 = 1 \mid X_2 = 0, X_3 = 0)$?

CHEAT SHEET:

Estimating p for iid Bernoulli:
$se(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
95% confidence interval for p: $\hat{p} \pm 2se(\hat{p})$.

Discrete Random Variable:
Given $E(X) = \sum x_i \, P(X = x_i)$.
$Var(X) = \sum (x_i - E(X))^2 \, P(X = x_i)$.
$sd(X) = \sqrt{Var(X)}$.

Joint distributions:
$p(x, y) = p(x) \, p(y|x)$,
X and Y are independent if $p(y|x) = p(y)$ for all x, or $p(x, y) = p(x)p(y)$.
Bayes: $p(x|y) = \frac{p(y|x)p(x)}{\sum p(y|x_i)p(x_i)}$.

Normal:
If $X \sim N(\mu, \sigma^2)$ then
$P(\mu - 2\sigma < X < \mu + 2\sigma) = .95$
$P(\mu - \sigma < X < \mu + \sigma) = .68$