MATH 1342

Review for Exam 2 (Alternate)

24 Multiple Choice Questions

Terms and Vocabulary:

Know sampling techniques

Experiment vs. Observational Study

Uniform Distribution vs. Normal Distribution

Normal Distribution (x) vs. Standard Normal Distribution (z)

Answer the following for a normal distribution with a mean of 65 and a standard deviation of 5.

P(X < 68)

P(X > 70)

P(57 < X < 60)

P(X < c) = 0.6753

P(X > c) = 0.3345

```
> pnorm(68,65,5)
[1] 0.7257469
> 1-pnorm(70,65,5)
[1] 0.1586553
> pnorm(60,65,5)-pnorm(57,65,5)
[1] 0.103856
> qnorm(0.6753,65,5)
[1] 67.27298
> qnorm(1-0.3345,65,5)
[1] 67.1376
```

Answer the following:

P(Z < 1.4)

P(Z > -0.8)

P(0.3 < Z < 1.9)

P(Z < c) = 0.7645

P(Z > c) = 0.4321

```
> pnorm(1.4)
[1] 0.9192433
> 1-pnorm(-0.8)
[1] 0.7881446
> pnorm(1.9)-pnorm(0.3)
[1] 0.353372
> qnorm(0.7645)
[1] 0.7208529
> qnorm(1-0.4321)
[1] 0.1710302
```

P(Z > c) = 1.4328 *No Answers (this would indicate a probability of larger than one, which is impossible)*

Random samples (of size 40) are selected from a population with μ = 345 and σ = 15. Determine the following:

Mean of the Sample Means ($\mu_{\bar{x}}$): $\mu_{\bar{x}} = \mu = 345$

Standard Error of the Sample Means ($\sigma_{\bar{x}}$): $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{40}} = 2.3717$

Probability that a randomly selected sample (as described above) would have a mean of greater than 370.

```
> 1-pnorm(370,345,15/sqrt(40))
[1] 0
>
```

Don't be scared of probabilities of zero!

A researcher wished to run an ***experiment*** concerning which color wall paint makes people more productive in a work environment. Determine the appropriate course of action:

❑ Send 100 people a survey concerning their work habits and the room color.

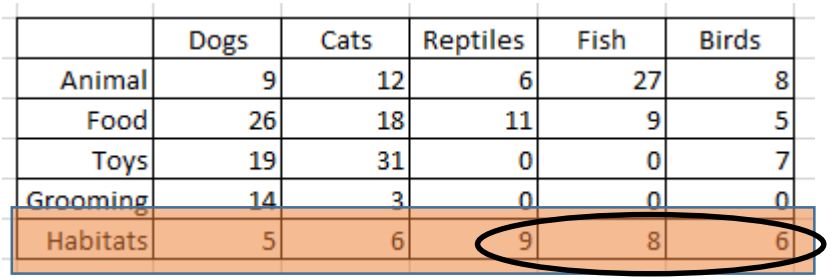
☑ Randomly assign 100 people to three different-colored room and monitor their results in a standardized skill-test. *Active control is applied here, therefore it is an experiment.*

❑ Observe the most productive employee in a corporation for a week to see what color room he/she usually works in.

❑ Speak to the head of human resources of a major corporation to discuss employee retention rates as a function of room color.

The following table shows the sales of a pet store for one week (delineated by type of pet and type of purchase). Use this to answer the following:

	Dogs	Cats	Reptiles	Fish	Birds
Animal	9	12	6	27	8
Food	26	18	11	9	5
Toys	19	31	0	0	7
Grooming	14	3	0	0	0
Habitats	5	6	9	8	6



	Dogs	Cats	Reptiles	Fish	Birds
Animal	9	12	6	27	8
Food	26	18	11	9	5
Toys	19	31	0	0	7
Grooming	14	3	0	0	0
Habitats	5	6	9	8	6

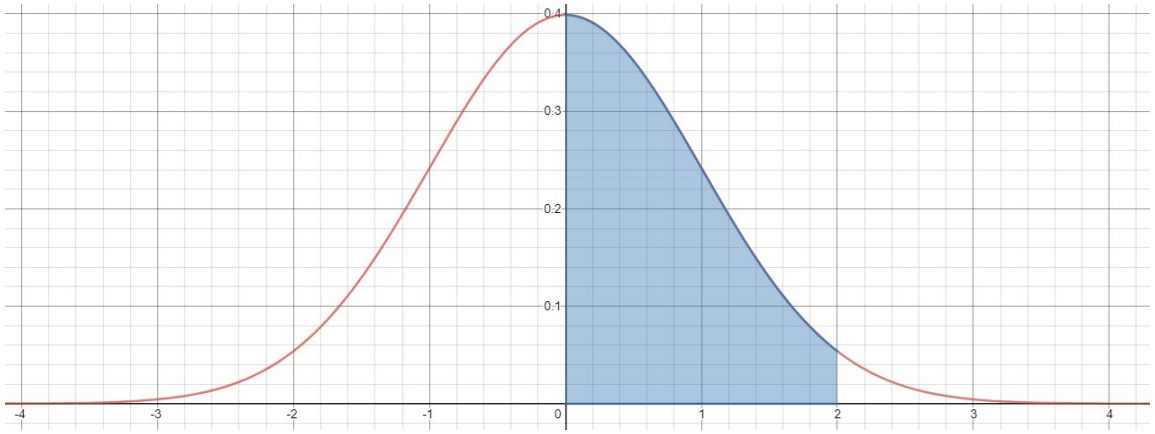
What is the probability that a dog owner purchased pet food?

$$P(Food|Dog) = \frac{26}{9 + 26 + 19 + 14 + 5} = \frac{26}{73} = 0.3561$$

If a customer was purchasing a habitat, what is the probability that their pet is not a dog or cat?

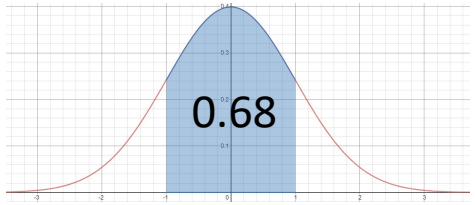
$$P((Dog \cup Cat)^c|Habitat) = \frac{9 + 8 + 6}{5 + 6 + 9 + 8 = 6} = \frac{23}{34} = 0.6765$$

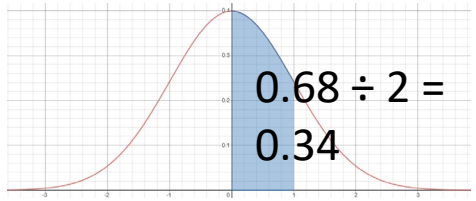
The following diagram represents continuous random variable of N(150, 35). The shaded region has an approximate area of ***0.475***. Determine the x-values of the left and right boundaries of the shaded region.

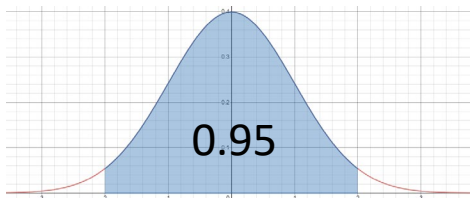


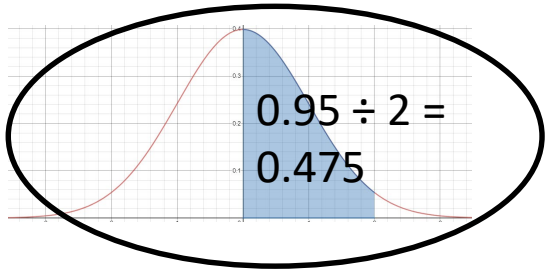
$[\mu , \mu + 2\sigma] = [150, 150 + 2 \cdot 35] = [150, 220]$

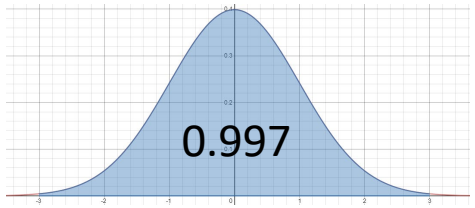
Using the Empirical Rule:

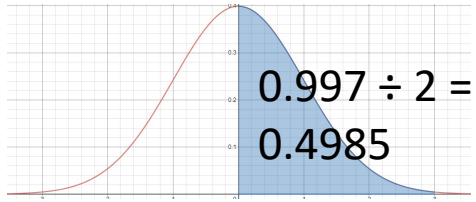












A regression analysis was done on a set of bivariate data. The table shows the resulting regression curve and its associated coefficient of determination. Rank the regression models from best fit to worst fit for this set of data.

Regression Model	R^2
$\hat{y} = -5.2x + 8.3$	0.8735
$\hat{y} = 0.43x^2 + 4.2x - 2.1$	0.7492
$\hat{y} = 4.53 + 6.22 \cdot \ln(x)$	0.4982
$\hat{y} = 1.543 \cdot e^{0.332x}$	0.9543

Regression Model	Ranking
Exponential	1 (Best Fit)
Linear	2
Logarithmic	4 (Worst Fit)
Quadratic	3

Ranking is based on r^2 values closest to 1 (Best Fit) through closest to 0 (Worst Fit).

Two males are comparing their heights. The first has a height of 71.3 inches and the second has a height of 178 centimeters. Which is taller? (Consider that mean height for a male, in inches, is 69.6986, standard deviation of 3.0023, and in centimeters is 176.9, standard deviation of 7.62.

$$Z_{inches} = \frac{x - \mu}{\sigma} = \frac{71.3 - 69.6986}{3.0023} = 0.5334$$

$$Z_{centimeters} = \frac{x - \mu}{\sigma} = \frac{178 - 176.9}{7.62} = 0.1444$$

The male with height of 71.3 inches is taller.

A cereal company has discovered that one of their coloring agents can cause a severe allergic reaction in 10% of the population, and a mild reaction in an additional 20% of the population. They decide to simulate an experiment to determine the severity of this problem. How would make assignments in a random digit table to model this situation?

Categories: Severe Reaction (10%); Mild Reaction (20%); No Reaction (70%)
Severe Reaction: 0 *(1 out of 10 digits: 10%)*
Mild Reaction: 1, 2 *(2 out of 10 digits: 20%)*
No Reaction: 3, 4, 5, 6, 7, 8, 9 *(7 out of 10 digits: 70%)*

A density curve consists of two line segments. The first connects the points (0, 1) and (0.5, 1). The second connects (0.5, 1) and the x-axis. *

Sketch the density curve.

Determine the coordinates where the graph crosses the x-axis

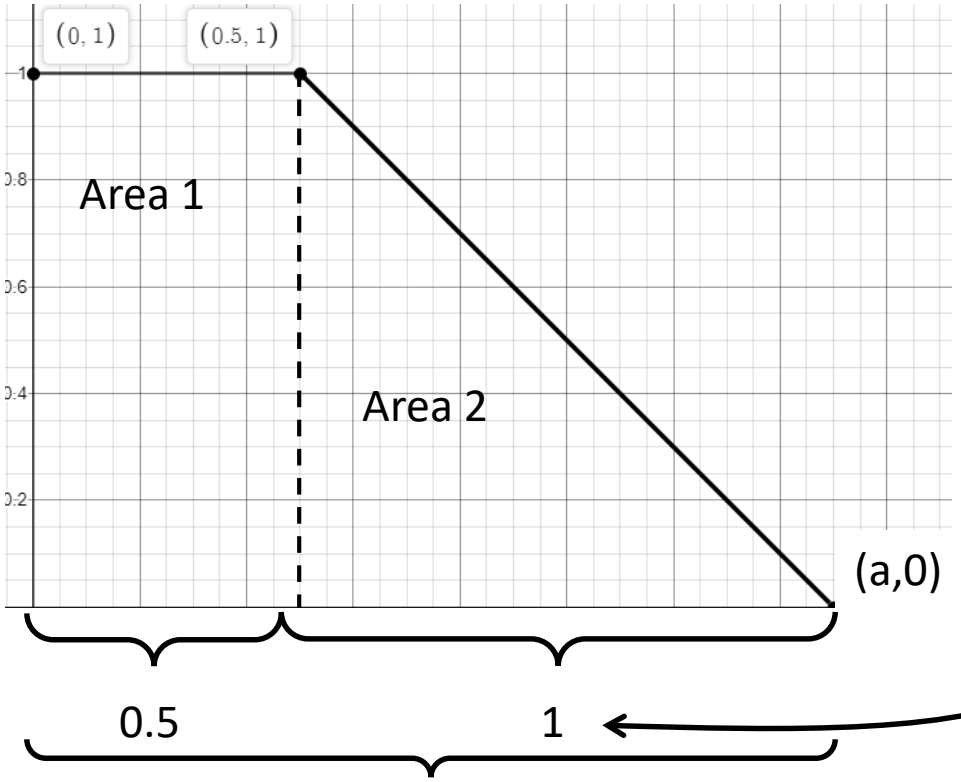
Calculate $P(x > 0.2)$

Calculate $P(x < 0.5)$

Calculate $P(x > 1.5)$

Determine the median

Continued:



Area 1 (Rectangle): $A = b \cdot h = 0.5 \cdot 1 = 0.5$
Area = Area 1 + Area 2
1 = 0.5 + Area 2
Area 2 =0.5

Area 2 (Triangle): $A = \frac{1}{2} \cdot b \cdot h \rightarrow 0.5 = \frac{1}{2} \cdot b \cdot 1 \rightarrow 0.5 = 0.5 \cdot b \rightarrow b = 1$

$P(X > 0.2) = P(0.2 < X < 0.5) + P(X > 0.5) = (0.3 \cdot 1) + \text{Area 2} = 0.3 + 0.5 = 0.8$ *(Method 1)*
$P(X > 0.2) = 1 - P(X < 0.2) = 1 - (0.2 \cdot 1) = 1 - 0.2 = 0.8$
(Method 2)

$P(X < 0.5) = \text{Area 1} = 0.5$
$P(X > 1.5) = 0$ *(Density Curve Only Exists for $0 < X < 1.5$)*
Median: $X = 0.5$ *(Since $P(X < 0.5) = P(X > 0.5) = 0.5$)*

The following table shows the heights (in inches) of several adolescents of different ages.

Age	Height
11	58.7
12	59.8
13	61.8
14	63.8
15	64.2
16	64.6
17	66.9

Identify the explanatory and response variables

Create a scatterplot of the data

Find the LSRL

Find and interpret r and r^2. Is this model a good fit for the data?

Use this model to extrapolate the height of a 20 year old.

Find the residual of a 15 year old.

Continued:

Explanatory: Age *("cause")*; Response: Height *("effect")*

```
> assign("age",c(11,12,13,14,15,16,17))
> assign("height",c(58.7,59.8,61.8,63.8,64.2,64.6,66.9))
> plot(age,height)
> regline=lm(height~age)
> regline

Call:
lm(formula = height ~ age)

Coefficients:
(Intercept)          age
     44.529        1.307
```

$$\hat{y} = 1.307x + 44.529$$

```
> cor(age,height)
[1] 0.9799978
> cor(age,height)^2
[1] 0.9603958
> 1.307*(20)+44.529
[1] 70.669
> 64.2-(1.307*15+44.529)
[1] 0.066
> residuals(regline)
          1           2           3           4           5           6
-0.20714286 -0.41428571  0.27857143  0.97142857  0.06428571 -0.84285714
          7
 0.15000000
```

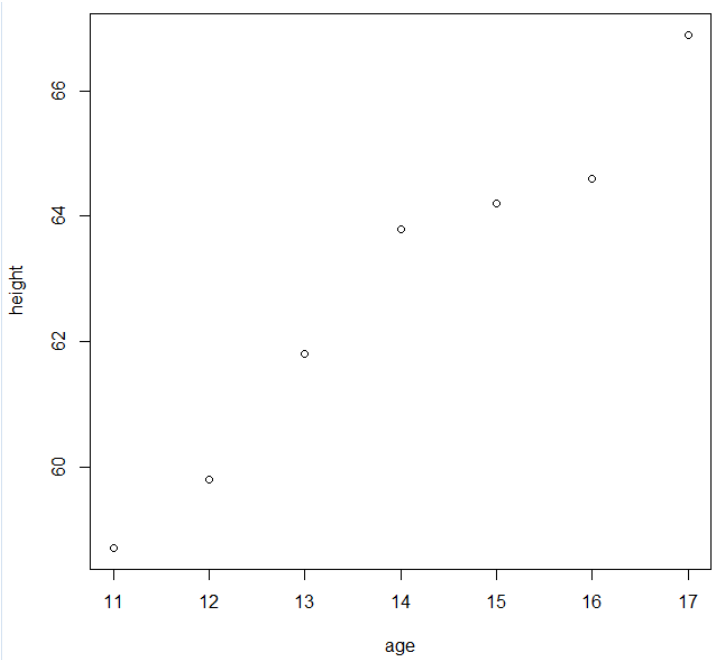
r: There is a strong positive relationship between the variables

r^2: 96% of the variation in height is explained by the LSRL

This model is a good fit to the data.

Method 1

Method 2



A researcher is running a study to see what kind of food purchases people make (generic versus name brand). In order to do this, he selects 100 people entering the supermarket and monitors them (through security cameras) to see which items they select.

Is this an experiment or an ***observational study***? Why? *There is no active control imposed by the researcher.*

Identify the population *(all people)* and sample *(100 selected people)*.

The researcher wants 25% of his sample to be senior citizens. How can probability sampling be used to do this? *When selecting a sample, make sure that 25% of the sampling frame are senior citizens. Then make a random selection.*

If the researcher also wants to consider race of the customer, what is this considered? *Blocking (introducing an additional factor that may influence the outcome)*

If the researcher removed all customers with children from the sample, what is this considered? *Under-Representation Bias*

*

A researcher is trying to determine if eating foods high in protein versus high in sugar has an immediate effect on motor skills and hand-eye coordination. To do this, he selects 45 people and divides them into three groups (2 treatment groups and a control group). After eating a meal, the subjects bowl five frames and the number of pins knocked down and the number of gutter balls is recorded.

Is this an ***experiment*** or an observational study? Why? *Active control is imposed by the researcher*

Identify the population (*all people*) and sample (*45 selected participants*).

Describe how matching can be used in the assignment of the groups. *Each of the three groups should contain people with good, moderate, and poor bowling skills.*

Describe possible control and treatment groups based on this research.

Control: No food provided (or food without high protein or sugar content).

Treatment 1: Food provided with high protein and low sugar.

Treatment 2: Food provided with high sugar and low protein.

Describe how double-blindness can be used in this research.

To reduce bias, the subjects will not know what is being testing in the experiment. The researcher will not know what food options subjects were provided until after all data was collected.