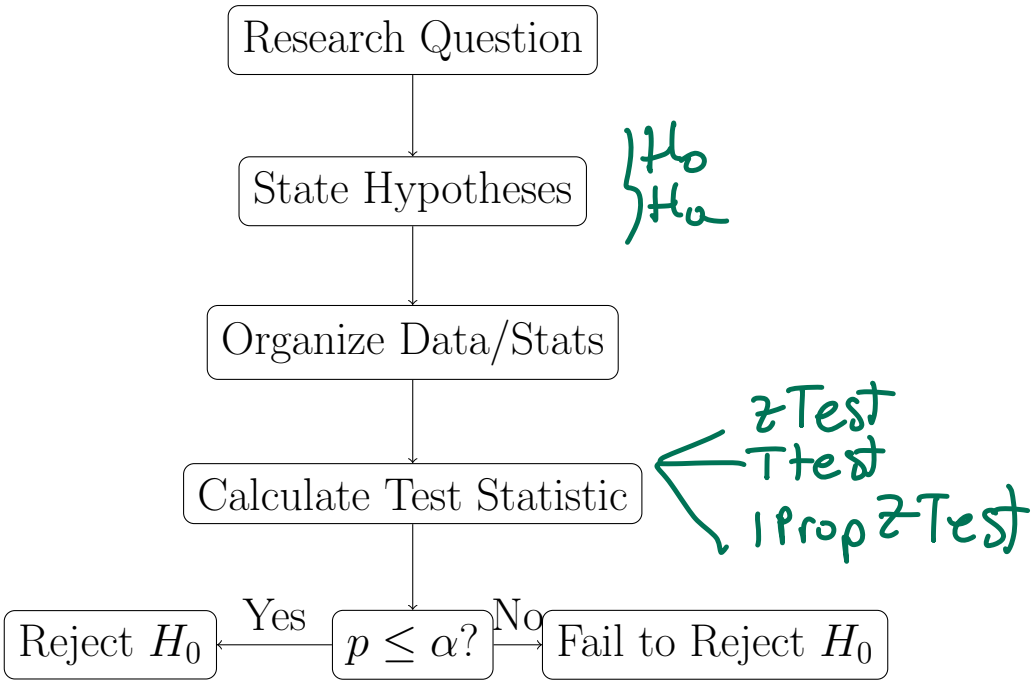
Math 142: Monday, Dec 2nd, Day 27 Notes

Agenda for Monday, Dec 2nd 2024

Coming up

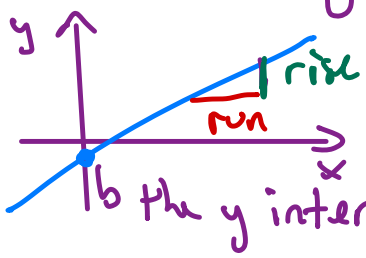
- Friday, Dec 6th Deadline for any late assignments (read Canvas announcement for more details)
- Excel 5 - optional (to replace a low Excel assignment grade)
- Final Exam **Wednesday, Dec 11th, 12PM-1:50PM in our classroom C168**

The Hypothesis Testing Process



Recall: Linear Equations

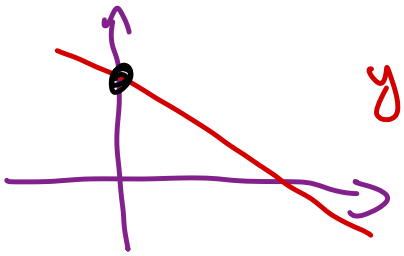
$y = mx + b$ in Algebra where m = slope and b was the y intercept



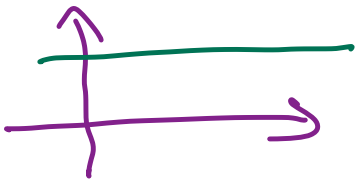
$y = mx + b$

m = slope $= \frac{\text{rise}}{\text{run}} = \frac{\text{change in } y}{\text{change in } x}$

↓ positive



$y = mx + b$ where slope is negative



$y = mx + b$ $\quad m = 0$

$y = b$

horizontal line

Example 1. Aaron's Word Processing Service (AWPS) does word processing. The rate for services is \$32 per hour plus a \$31.50 one-time charge. The total cost to a customer depends on the number of hours it takes to complete the job. Find the equation that expresses the total cost in terms of the number of hours required to complete the job.

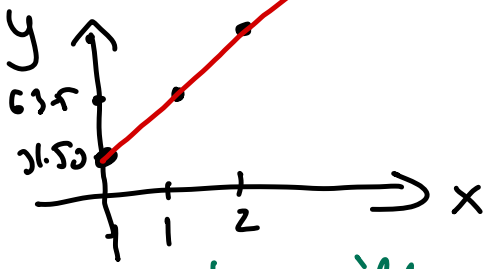
$y = mx + b$

x: input; independent variable: # of hours

y: output: dependent variable: total cost

$y = mx + b$

$y = 32x + 31.50$



x	y
1	$32 + 31.50 = 63.50$
2	$32(2) + 31.50 = 95.50$
⋮	⋮

How much will a job that take 5 hours cost?

$y = 32(5) + 31.50 = \$191.50$

Example 2. House Prices Imagine you're a real estate agent trying to estimate house prices in your area. You notice that the size of a house (in square feet) seems to influence its price. Let's look at some data:

Size (sq ft)	Price ($1000s)
1200	240
1500	285
1800	330
2100	~~365~~ 400
2400	405

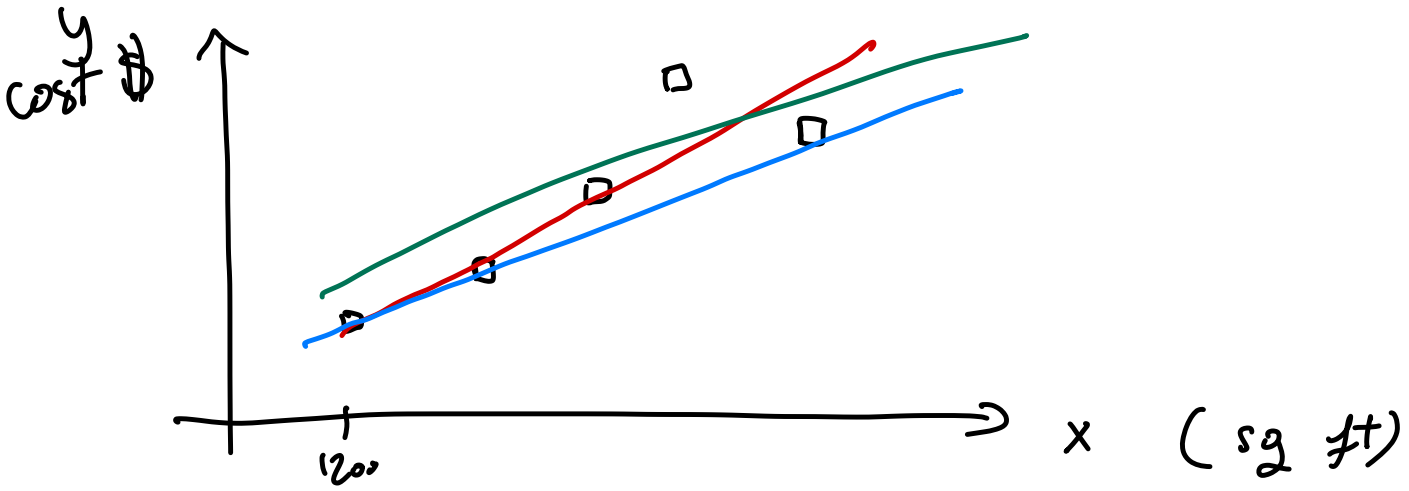
Questions that arise:

- Is there a relationship between house size and price?
- How can we quantify this relationship?
- Can we predict the price of a house given its size?
- What if I want to predict the price for a 1950 sq ft house?

Linear regression helps us answer these questions!

Scatterplot

A scatterplot is graph on the coordinate plane where there is one point for each pair of data (independent variable & dependent variable)



Understanding Linear Regression

Linear regression models the relationship between two variables:

- Independent variable (x): The predictor (e.g., house size)
- Dependent variable (y): The response (e.g., house price)

The Linear Model

The line that best fits the data is called the **line of best fit** or the **regression line**. The equation for a linear regression line is:

$$y = ax + b$$

where:

- y is the dependent variable (the predicted value)
- a is the slope
- x is the independent variable (the predictor value)
- b is the y intercept (the y value when $x = 0$)

We use the calculator to get the equation of the regression line and then we use the equation for predictions.

Caution: Interpolation vs Extrapolation!

Do not use the line of best fit to predict for x values outside of the available data.

For the real estate example we got the equation of the regression line to be $y = 0.148x + 65$

Only use this model for x values between 1200 - 2400 (interpolation).

Using for values below 1200 or above 2400 is called extrapolation.

Step-by-Step Calculator Instructions

Calculator Instructions

1. Enter Data:
   - Press **STAT** then Select **1:Edit...**
   - Enter x-values in L1 and y-values in L2
2. Create Scatter Plot:
   - Press **2nd** → **Y=** (STAT PLOT)
   - Select **Plot1** and Turn plot **On**
   - Select scatter plot type (first option)
   - Set Xlist: L1, Ylist: L2
   - leave FregList blank
   - In Store RegEq: press **Vars**, then **Y-Vars**, pick **1:Function** and then select Y_1
   - Press **ZOOM** → **9:ZoomStat**
3. Calculate Regression:
   - Press **STAT** → **CALC**
   - Select **4:LinReg(ax+b)**
   - Press **ENTER**
4. View Results:
   - a = slope
   - b = y-intercept
   - r = correlation coefficient
   - r^2 = coefficient of determination

$y = ax + b$

$y = 0.148 \quad b = 65$

$y = 0.148x + 65$

$y = 0.148(1950) + 65$

$y = 353.6$

$y = 353\,600$

What is Correlation?

The correlation coefficient (r) tells us:

- Direction of relationship (positive or negative)
- Does not imply cause and effect. Only indicates that variables tend to move together
- Strength of relationship (how close to ± 1)
- Perfect relationship: $r = \pm 1$ and No linear relationship: $r = 0$

Interpreting Correlation Values

$|r| \leq 0.8$ means $-0.8 \leq r \leq 0.8$

- $|r| \leq 0.8$: Strong correlation
- $0.5 \leq |r| < 0.8$: Moderate correlation
- $0.2 \leq |r| < 0.5$: Weak correlation
- $|r| < 0.2$: Very weak or no correlation

Correlation Visualizations

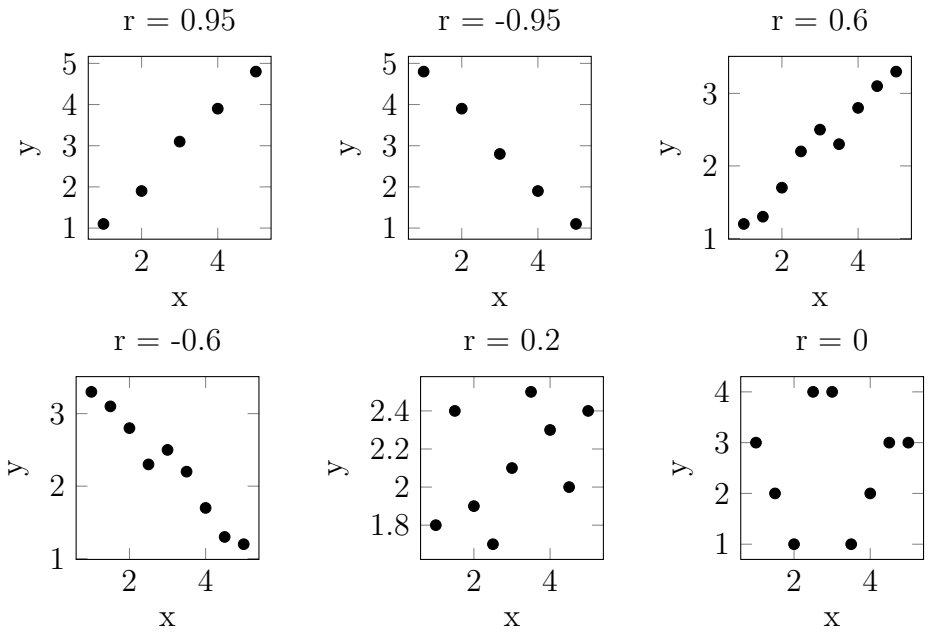


Figure 1: Examples of Different Correlation Values