Explained variation – amount of variation in a data set that a mathematical model can account for
Unexplained variation – amount of variation that the model cannot explain

Decision Errors

Type 1 Error – reject the null hypothesis when H0 is actually true.
Type 2 Error – failing to reject the null hypothesis when the alternative is actually true.

Example:

The actual truth

		H0 True	H0 False	
Experiment prediction	Reject H0	60	120	180
	Fail to Reject H0	40	80	120
		100	200	

Type 1 error = P(Reject H0 | H0 True) = 60/100
Type 2 error = P(Fail to Reject H0 | H0 False) = 80/200

Tests

Binomial test

Test for the proportion of success.

binom.test(x = 31, n = 39, p = 0.7, conf.level = 0.8)

- x – the number of “successes” we observed
- n – the total number of observations
- p – our hypothesized proportion

t-test

Used when standard deviation is **unknown** and the sample size is **large**.

One sample t-test:	Compare the mean of a single sample to a known value Null hypothesis: $\mu = \mu_0$ (hypothesized mean is the same as true mean)	$t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$

	df = n - 1	$\overline{X}$ = sample mean μ_0 = Hypothesized population mean s = sample standard deviation n = sample size
Two sample t-test:	Compare the means of two independent groups to determine if there is a significant difference between them. Null hypothesis: $\mu_1 = \mu_2$ $df = n_1 + n_2 - 2$	$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ $\overline{x}_1, \overline{x}_2$ = means of the two samples s_1^2, s_2^2 = variances of the two samples n_1, n_2 = sizes of the two samples
Difference in proportion test:	Compares the means of two related groups, such as before-and-after measurements on the same subjects. df = n - 1	

Steps to Perform a T-Test

1. State Null hypothesis and Alternative hypothesis.
2. Calculate the Test Statistic
3. Determine Degrees of Freedom
4. Find the Critical Value:
    - Use a t-table for the given df and significance level(α).
5. Access the significance of hypothesis test:
    - Compare t-Value to Critical Value – If t-value exceeds the critical value, reject H0.
    - Compare p-value to significance level – If p-value smaller or equal to α, reject H0.

z-test

Used when the standard deviation is **known** and the sample size is **small**.

Chi Square Test

– Goodness-of-fit Tests (1 Variable)

$$Expected\ frequency = \frac{Row\ total \times Column\ total}{Grand\ total}$$

Null Hypothesis: the observed data follows the expected distribution.

Example:

Actual:

	Category A	Category B	Sum
Group 1	20	30	50
Group 2	40	10	50
Sum	60	40	100

Expected:

	Category A	Category B
Group 1	30	30
Group 2	20	20

Group 1 & Category A:
$$Expected = \frac{60 \times 50}{100} = 30$$

Group 1 & Category B:
$$Expected = \frac{50 \times 40}{100} = 20$$

Group 2 & Category A:
$$Expected = \frac{50 \times 60}{100} = 30$$

Group 2 & Category B:
$$Expected = \frac{50 \times 40}{100} = 20$$

– Independence Test (2 Variable)

$$X^2 = \sum \frac{(Observed\ Value - Expected\ Value)^2}{Expected\ Value}$$

Null Hypothesis: there is no relationship between the two variables.

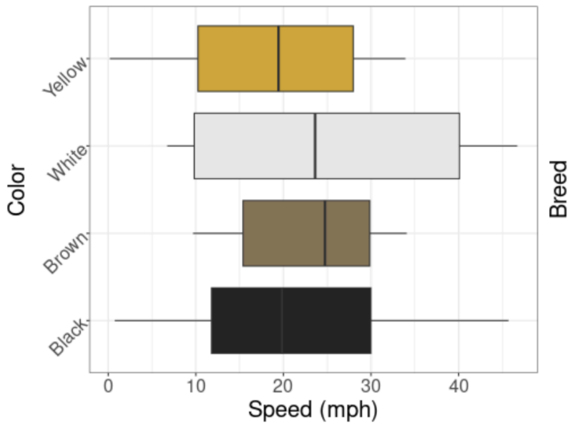
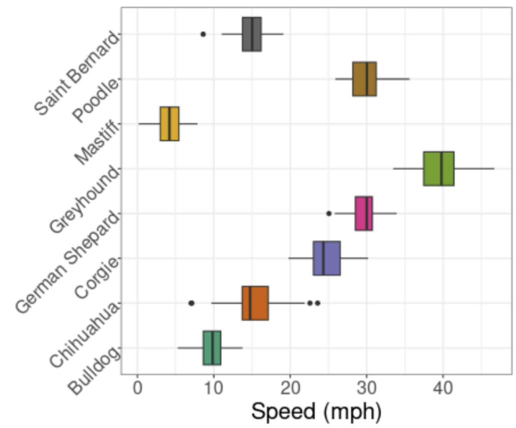
ANOVA (Analysis of Variance)

a collection of statistical models used to analyze difference among many means
ANOVA only tells us that a difference exists or not, not where it is or to what degree

Conditions need to meet before conducting ANOVA:

- Observations are independent within and between groups,
- Responses within each group are nearly normal
- Variability across the groups is about equal.

Null hypothesis: there are no difference between the mean for all groups ($\mu 1 = \mu 2 = \cdots = \mu k$).

 Smaller F value	 Greater F value
Low variability between groups High variability within groups	High variability between groups Low variability within groups

$$\underbrace{\sum_{i}^{n}(x_{ij}-\overline{x})^2}_{\text{Total variability}} = \underbrace{\sum_{i=1}^{n}(x_{ij}-\overline{x}_j)^2}_{\text{Variability within groups}} + \underbrace{\sum_{j=1}^{k} n_j(\overline{x}_j-\overline{x})^2}_{\text{Variability between groups}}$$

Total variability = SSE (Variability within groups) + SSG (Variability between groups)

Sum of square error

$$SSG = \sum_{i=1}^{n} (x_{ij} - \overline{x}_j)^2$$

x_{ij} = observed value

$\overline{x}_j$ = predicted value

Sum of Squares error between Groups

$$SSG = \sum_{j=1}^{k} n_j(\overline{y}_j - \overline{y})^2$$

n_j = number of observations in group j

$\overline{y}_j$ = mean of the response variable for group j

$\overline{y}$ = total mean of the response variable

Mean squared error

$$MSE = \frac{SSE}{n-k}$$

k = number of groups

n - k = degree of freedom

Mean squared error between groups

$$MSG = \frac{SSG}{k-1}$$

k = number of groups

$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{Y}_i)^2$ n = number of data points Y_i = observed value $\widehat{Y}_i$ = predicted value	
Standard Error $SE = \sqrt{\frac{sum\ of\ square\ error}{n}}$	**F-value** – can be used as a measure of certainty. Greater F-value is better. Large F value: the differences in group means are greater random variation from the differences within the groups. Small F value: the differences in means are likely caused by the differences within the groups. $F = \frac{MSG}{MSE}$

Why use ANOVA instead of t-test:

1) Need to recalculate the significant value. (family-wise error rate)
2) Time consuming

Post-hoc Test

Analyze how much variation are between each pair of groups

Use the Tukey Range Test or the Tukey Honest Significant Difference Test

Type of Test		Type of variable	Description
t-test	Difference in mean	One categorical One quantitative	Compare the mean of two independent groups Same as anova, but needs to adjust family wise error (Bonferonni correction) for more than two groups.
	Paired sample	One categorical One quantitative	Compare the mean of two related groups (ex. before and after)
	One sample	One quantitative	Compare the actual mean to the predicted mean

	Difference in proportion	Two categorical (success/fail, yes/no)	Compare the proportion of two independent groups Same as independence chi square test
Chi square test	Goodness of fit	One categorical	Compare the actual number of count to the predicted number of count in each category
	Independence	Two categorical (success/fail, yes/no)	
Anova		One quantitative One categorical(with multiple groups within it)	Compare whether the mean of two groups is the same. Same as the difference in mean t-test.
Tukey test		One quantitative One categorical (with multiple groups within it)	Compare how much difference there is between the means of groups.
Linear Regression		Anything	