Chapter Nine

Hypothesis Testing with r, χ^2, and F

Relationships Between Two Continuous Variables: Linear Correlation

Relationships Between Two Discrete Variables: Independence

Distributions: Goodness of Fit

Standard Deviations: Single Variance

Differences Between Standard Deviations: Two Variances

Differences Between Means: ANOVA

Selecting a Statistical Test

Statistical tests

Previously, we only tested one or two means or proportions. However, there are many testable scenarios beyond these.

Parameters tested	Statistic	Alternate Hypothesis
Single Mean	t or z	The mean is higher than a stated value.
Single Proportion	z	The proportion is higher than a stated value.
Single Mean Difference	t or z	One mean is higher than the other (within-participants).
Two Means	t or z	One mean is higher than the other (between-participants).
Two Proportions	z	One proportion is higher than the other.
Correlation	r	The two numerical variables are correlated.
Independence	χ^2	The two nonnumerical variables are not independent.
Distribution	χ^2	The distribution is different than a stated distribution.
Single Variance	χ^2	The standard deviation is higher than a stated value.
Two Variances	F	One standard deviation is higher than the other.
Multiple Means	F	Not all of the means are equal.

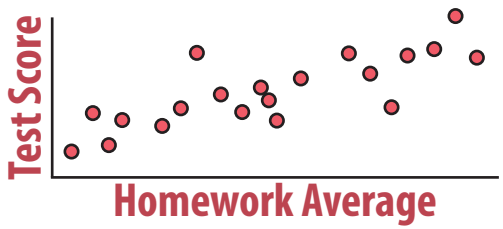
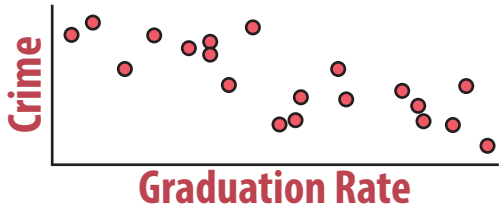
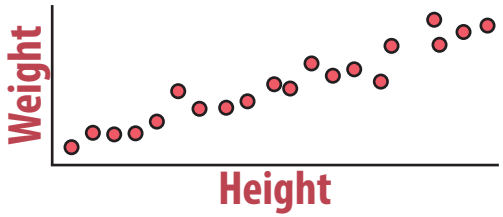
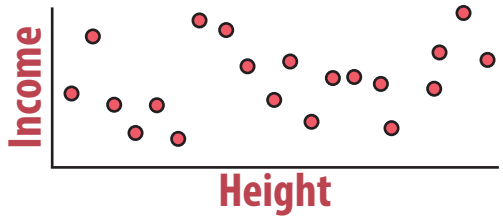
Doing a statistical test

Each type of statistical test is done differently, but the general approach is the same for all.

Step	By hand	On calculator
Choose the tail(s)	Identify the tail(s) of the test before you know the sample statistics or data.	Identify the tail(s) of the test before you know the sample statistics or data.
Enter the data or stats	Put the data or statistics into tables or formulas as specified.	Enter the data into lists, or enter the statistics directly.
Set up the test	Look up the critical value in the appropriate table.	Highlight $\neq$ for two-tailed, $>$ for right tailed, or $<$ for left-tailed.
Make the statistical conclusion	Reject H_0 if the calculated statistic is in the critical region.	Reject H_0 if p is less than .05.

The *r* distribution

The **correlation coefficient *r*** is a value between -1 and 1 that summarizes the strength and direction of the relationship between the two variables in the sample.

Value of *r*	Correlation	Meaning	Example
Positive	positive	The higher *x* is, the higher *y* tends to be.	
Negative	negative	The higher *x* is, the lower *y* tends to be.	
Close to 1 or -1	strong	The two variables are closely related in the sample, so the line of best fit is a good predictor.	
Close to 0	weak	The two variables are not closely related in the sample, so the line of best fit is a poor predictor.	

Calculating *r*

See the notes for details and examples.

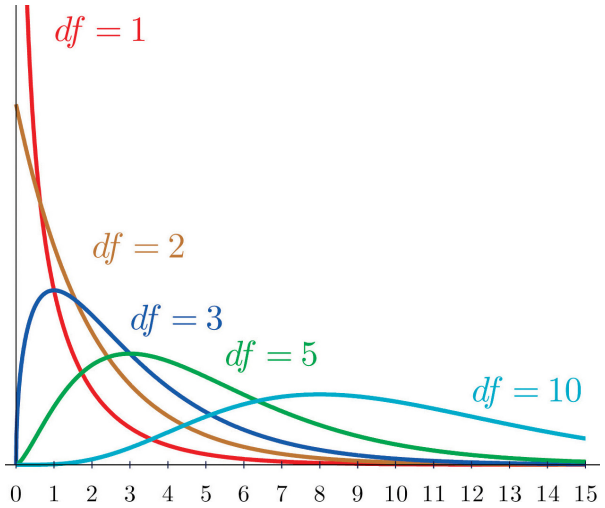
Test	By hand	By calculator
Correlation	Multiply each $x - \bar{x}$ by the corresponding $y - \bar{y}$, find the sum of these products, and divide by $n - 1$. Divide this quotient, called the covariance, by both standard deviations.	Enter the data into two lists, choose `LinRegTTest`, and set the tails.

The χ^2 distribution

The **chi-square** distribution with one degree of freedom is the distribution of squared z scores: $\chi^2 = z^2$. For $df > 1$, χ^2 is the sum of df squared z scores.

The χ^2 distribution is similar to the normal curve but is not symmetrical. Instead, it starts at $\chi^2 = 0$ and peaks at $\chi^2 = df - 2$.

The χ^2 table shows the critical value needed to have the stated area to the right. For right-tailed tests, this area is .05; for left-tailed it is .95, and for two-tailed it is .025 on the right and .975 on the left.



χ^2 test	Tails	Degrees of Freedom	Example
Independence	right	$(k_1 - 1)(k_2 - 1)$	Is gender independent from season of birth? $df = (2 - 1)(4 - 1) = 3$
Goodness of Fit	right	$k - 1$	Is each side of a 6-sided die equally likely to be rolled? $df = 6 - 1 = 5$
Single Variance	right, left, or two	$n - 1$	Based on a sample of 20 IQ scores, is σ higher than 15? $df = 20 - 1 = 19$

Calculating χ^2

See the notes for details and examples. Not all of these tests can be done on all calculators.

Test	By hand	By calculator
Independence	Count the number of data (O) in each condition, calculate (proportionally) the expected number of data (E) in each condition, and use these values to calculate $\chi^2 = \sum \frac{(O-E)^2}{E}$.	Count the number of data in each condition, enter these counts into a matrix, and choose χ^2`-Test`.
Goodness of Fit	same as above	Count the number of data (O) in each condition and list these in `L1`, and calculate (proportionally) the expected number of data (E) in each condition and list these in `L2`. Then choose χ^2`GOF-Test` and enter $k-1$ for `df`.
Single Variance	Divide the sample variance by the population variance being tested, and multiply by degrees of freedom: $\chi^2 = (n-1)\frac{s^2}{\sigma^2}$.	n/a

The *F* distribution

The ***F*** distribution is a ratio of variances: $F = \frac{s_1^2}{s_2^2}$. As defined earlier, variance is sum of squares divided by degrees of freedom. Given that χ^2 is a sum of squares, *F* can be defined as $F = \frac{\chi_1^2 / df_1}{\chi_2^2 / df_1}$. Note that *F* has different degrees of freedom for the numerator than the denominator.

The most common use of *F* is for analysis of variance, or ANOVA. For ANOVA, the numerator variance is the variance between the samples, and the denominator variance is the variance within the samples.

The *F* distributions starts at zero. If the numerator has more than one degree of freedom, the peak of the distribution is between $F = 0$ and $F = 1$. Otherwise, the peak is at 0, in which case $F = t^2$.

F test	Tails	Degrees of Freedom	Example
Two Variances	right or two	$df_N = n_1 - 1$ $df_D = n_2 - 1$	Based on a sample of 9 kids and 6 adults, does kids' sleep vary more? $df_N = 9 - 1 = 8, df_D = 6 - 1 = 5$
ANOVA	right	$df_N = k - 1$ $df_D = (n_1 - 1) + (n_2 - 1) + \ldots$	Based on a sample of 9 kids, 6 adults, and 5 babies, does sleep vary by age? $df_N = 3 - 1 = 2$ $df_D = (9 - 1) + (6 - 1) + (5 - 1) = 17$

Calculating *F*

See the notes for details and examples. Not all of these tests can be done on all calculators.

Test	By hand	By calculator
Two Variances	Divide the variance that was predicted to be larger by the variance of the other sample. (For two tailed, use the actual larger variance as the predicted larger variance.)	If the sample variances are already calculated, choose `2-SampFTest`, select `stats`, and set the tails. Otherwise, enter the data into two lists, choose `2-SampFTest`, select `data`, and set the tails.
ANOVA	Find *SS* and *df* for each sample, and divide the total of these sums of squares by the total of these degrees of freedom to get the variance within samples. Use weighted averages to find the sum of squares between samples, and divide this by the degrees of freedom between samples to get the variance between the samples. Divide the between-sample variance by the within-sample variance.	Enter each data set into a separate list, choose `ANOVA`, and type the list names, separated by commas (e.g., `ANOVA(L1,L2,L3)`).

Distinguishing between statistical tests

Test	What to look for	Example
Single Mean	one sample, with averaged values	Do students do more than 5 hours of homework per week?
Single Proportion	one sample, for which a portion of the data meet the criterion	Do more than 80% of students do more than 5 hours of homework per week?
Single Mean Difference	a within-participants design	Do students spend more time on homework their junior year than their senior year?
Two Means	a between-participants design	Do girls spend more time on homework than boys?
Two Proportions	two samples, for which a portion of the data meet the criterion	Are girls more likely than boys to complete their homework?
Correlation	two numerical variables	Does amount of time on homework freshman year predict amount of time on homework junior year?
Independence	two nonnumerical variables	Does science class taken vary by sex?
Goodness of Fit	a distribution	Is the amount of homework done per week normally distributed?
Single Variance	one sample, testing variability (or consistency)	Is the standard deviation of homework hours per week greater than 2 hours?
Two Variances	two samples, comparing variability (or consistency)	Do boys vary more than girls on hours of homework per week?
ANOVA	more than two samples	Does hours of homework vary by grade level?