HYPOTHESIS TESTING

four steps
1. State the hypothesis and the criterion
2. Compute the test statistic.
3. Compute the $p$-value.
4. Make a decision

HYPOTHESIS TESTING

we need to know the properties of the sampling distribution
for the mean, the central limit theorem tells us that the sampling distribution is normal, and specifies the mean and standard deviation (standard error)
area under the curve of the sampling distribution gives probability of getting that sampled value, or values more extreme ($p$-value)
for other types of statistics, the sampling distribution is different
area under the curve of sampling distribution still gives probability of getting that sampled value, or values more extreme
correlation coefficient

CORRELATION COEFFICIENT

from a population with scores $X$ and $Y$, we can calculate a correlation coefficient
let $\rho$ be the correlation coefficient
$\text{parameter}$ of the population
let $r$ be the correlation coefficient
$\text{statistic}$ from a random sample of the population

SAMPLING

depending on which points we sample, the computed $r$ will take different values
RANDOM SAMPLING

 Sampling Distribution

 frequency of different r values, given a
 population parameter \( \rho \)

 not usually a normal distribution!
 often skewed to the left or the right
 cannot find area under curve!

 FISHER z TRANSFORM

 formula for creating new statistic

 \[
 z_r = \frac{1}{2} \log_e \left( \frac{1 + r}{1 - r} \right)
 \]

 where

 \( \log_e \) is the “natural logarithm” function
 also sometimes designated as \( \ln \)

textbook provides a r to z’ calculator

 FISHER z TRANSFORM

 means we can use all our knowledge
 about hypothesis testing with normal
 distributions for the transformed
 scores!

 online calculator converts r to zr (it
 calls it z’)

e.g.

 \[
 r = -0.90 \rightarrow z_r = -1.472 \\
 r = 0 \rightarrow z_r = 0 \\
 r = 0.45 \rightarrow z_r = 0.485
\]

 we can convert back the other way
 from zr \( \rightarrow \) r too!

 HYPOTHESIS TESTING

 Suppose we study a population of data
 that we think has a correlation of 0.65.
 We want to test the hypothesis with a
 sample size of \( n = 30 \).
 (e.g. family income and attitudes
 about democratic childrearing)

 Step 1. State the hypothesis and
 criterion

 \( H_0 : \rho = 0.65 \)
 \( H_A : \rho \neq 0.65 \)

 two-tailed test

 \( \alpha = 0.05 \)
HYPOTHESIS TESTING

Step 2. Compute the test statistics

Suppose from our sampled data we get

\[ r = 0.61 \]

we need to convert it to a z-score

\[ z = r - z_{p} = 0.61 - 0.709 \]

and calculate standard error

\[ s_{z_{r}} = \frac{1}{\sqrt{n - 3}} = \frac{1}{\sqrt{27}} = 0.192 \]

HYPOTHESIS TESTING

Step 2. Compute the test statistics

now we calculate the test statistic

\[ Test\ statistic = \frac{statistic - parameter}{standard\ error\ of\ the\ statistic} \]

\[ z = \frac{r - \mu}{s_{z_{r}}} = \frac{0.700 - 0.775}{0.192} = -0.344 \]

Step 3. Compute the p-value

From the Normal Distribution calculator, we compute

\[ p = 0.7346 > 0.10 = \alpha \]

HYPOTHESIS TESTING

Step 4. Make a decision.

\[ p = 0.7346 > 0.05 = \alpha \]

H_{0} \text{ is not rejected at the 0.05 significance level} \]

The probability of getting \( r = 0.61 \) (or a value further away from 0) with a random sample, if \( \rho = 0.65 \), is greater than 0.05.

The observed difference is not a significant difference.

A SPECIAL CASE

hypothesis testing of correlation coefficients \( \rho \) and \( r \)
Fisher’s z transform

\[ H_{0}: \rho = a \]
\[ H_{a}: \rho \neq a \]

special case \( a = 0 \):

\[ H_{0}: \rho = 0 \]
\[ H_{a}: \rho \neq 0 \]

Is there a significant correlation coefficient?

SAMPLING DISTRIBUTION

while we needed Fisher’s z transform to convert the sampling distribution into a normal distribution

it is not necessary for testing \( \rho = 0 \)

for \( \rho = 0 \)

the sampling distribution of the test statistic is a t distribution with df = \( n - 2 \) (two sets of scores, minus 1 from each set)

no need to convert with Fisher z transform

we follow the same procedures as before

1. State the hypothesis. \( H_{0}: \rho = 0 \) and set the criterion
2. Compute the test statistic.
3. Compute the p-value
4. Make a decision.
HYPOTHESIS TESTING

everything is the same, except the test statistic calculation is a bit different

it turns out that an estimate of the standard error is:

\[ s_r = \sqrt{\frac{1 - r^2}{n - 2}} \]

so that the test statistic is:

\[ t = \frac{r - \rho}{s_r} = r \frac{n - 2}{\sqrt{n - 2}} \]

we use this with a t distribution to compute a p-value

EXAMPLE

\[ n = 32 \text{ scores calculated to get} \]
\[ r = -0.375 \]

1. State the hypothesis.
   \[ H_0 : \rho = 0, \ H_a : \rho \neq 0 \]
2. Set the criterion for rejecting \( H_0 \).
   \( \alpha = 0.05 \)
3. Compute the test statistics.
   \[ t = r \frac{n - 2}{\sqrt{1 - r^2}} = (-0.375) \frac{30}{0.859} = -2.216 \]
   use the t Distribution calculator with df = n - 2 = 30
   \[ p = 0.0344 \]
4. Interpret the results: \( p = 0.0344 < 0.05 = \alpha \) reject \( H_0 \)

EXAMPLE

I took the percentage of the first six homework grades and correlated it with the first exam scores
\[ r = -0.0584 \]
Is this a significant correlation?

CAREFULL!

If I treat the class as a population, the correlation is what it is. Significance is not an issue!

If I treat the class as a sample of students who do homework and take exams in statistics, then I can ask about statistical significance

EXAMPLE

For Homework and Reading, \( r = 0.836 \). I have \( n = 28 \) scores

Compute the test statistics.
\[ t = r \frac{n - 2}{\sqrt{1 - r^2}} = -0.296 \]
use the t Distribution calculator with df = n - 2 = 26
\[ p = 0.769 \]
Interpret the results:
\( p = 0.769 > 0.05 = \alpha \), do not reject \( H_0 \)

READING?

For Homework and Reading, \( r = 0.836 \). I have \( n = 28 \) scores

Compute the test statistics.
\[ t = r \frac{n - 2}{\sqrt{1 - r^2}} = 7.77 \]
use the t Distribution calculator with df = n - 2 = 26
\[ p \approx 0 \]
Interpret the results:
\( p \approx 0 < 0.05 = \alpha \), reject \( H_0 \)
CAREFUL!

When we conclude a test is statistically significant, we base that on the observation that observed data (or more extreme) would be rare if the $H_0$ were true.

But if we make multiple tests from a single sample, our calculations of probability may be invalid.

We performed two hypothesis tests from one sample of students.

Each test has a chance of producing a significant result, even if $H_0$ is true.

It is not appropriate to just run various tests with one data set, if all you are doing is looking for significant results (fishing).

You have to do a different type of statistical analysis.

CONFIDENCE INTERVAL

Always use the Fisher z transform.

Build interval as a Fisher z score and then convert to correlation ($r$ value).

$CI = z_r \pm z_{cv} \cdot s_{z_r}$

For the correlation between homework and reading scores:

$CI_{95} = 1.2084 \pm (1.96)(0.2) = (0.816, 1.600)$

when we convert to $r$ values:

$(0.673, 0.922)$

POWER

How would we design a good experiment to test a correlation?

How big a sample do we need to have a 90% chance of rejecting the null hypothesis?

Conceptually, this is the same issue as estimating power or sample size for a hypothesis test of means.

We just need to use the sampling distribution for the Fisher z transform of the sample correlation instead of the sampling distribution for a sample mean.

CONCLUSIONS

correlation coefficient

Fisher z transform

testing significance of correlation

confidence interval

power

Higher than 99.9% chance of rejecting the null hypothesis.

What sample size do we need to have 90% power?

However, whether these calculations make sense depends on whether $\rho = 0.8$ in reality.
NEXT TIME

hypothesis testing of two means

Why do we let people die?