HYPOTHESIS TESTING

four steps
1. State the hypothesis.
2. Set the criterion for rejecting.
3. Compute the test statistics.
4. Interpret the results.

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

\[
\rho = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y}
\]

let $\rho$ be the correlation coefficient $\rho$ parameter of the population (avoid confusion with Spearman’s $\rho$)

let $r$ be the correlation coefficient $r$ statistic from a random sample of the population

SAMPLING

$\rho = 0.22$

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

\[ r = 0.24 \]

\[ r = 0.67 \]

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

for large samples, the sampling distribution of \( z_r \) is normally distributed

(regardless of the value of \( \rho \))

with a mean

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

and with standard error (standard deviation of the sampling distribution)

\[ s_{z_r} = \frac{1}{\sqrt{n - 3}} \]

where \( n \) is the sample size

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.

\( H_0 : \rho = 0.65 \)

\( H_a : \rho \neq 0.65 \)

two-tailed test

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 \( z_r \) calculator

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 \( z_r \rightarrow r \) too!
HYPOTHESIS TESTING

Step 2. Set the criterion for rejecting $H_0$

$\alpha = 0.10$

Step 3. Compute the test statistics

Suppose from our sampled data we get $r = 0.61$

We need to convert it to a $z_r$ score

$r = 0.61 \rightarrow z_r = 0.709$

And calculate standard error

$s_{z_r} = \frac{1}{\sqrt{n-3}} = \frac{1}{\sqrt{27}} = 0.192$

now we calculate the test statistic

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

$z = \frac{z_r - z_\rho}{s_{z_r}} = \frac{0.709 - 0.775}{0.192} = -0.344$

From the Normal Distribution calculator, we compute

$p = 0.7346 > 0.10 = \alpha$

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 $z$ transformation to convert the sampling distribution into a normal distribution, it is not necessary for testing $\rho = 0$

SAMPLING DISTRIBUTION

For $\rho = 0$

the sampling distribution is a $t$ distribution with $df = n - 2$

(two sets of scores, minus 1 from each set)

No need to convert with $z$ transform

We follow the same procedures as before

1. State the hypothesis. $H_0 : \rho = 0$

2. Set the criterion for rejecting $H_0$

3. Compute the test statistics.

4. Interpret the results.
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:
\[ s_r = \frac{1 - r^2}{\sqrt{n - 2}} \]
so that the test statistic is:
\[ t = \frac{r - \rho}{s_r} = r \frac{n - 2}{\sqrt{1 - r^2}} \]
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.
   \[ t = r \frac{n - 2}{\sqrt{1 - r^2}} = \left( -0.375 \right) \frac{30}{0.839} = -2.216 \]
   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
\[ \rho = 0.288 \]
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

CAREFULL!

is \( r = 0.288 \) significantly different from \( 0 \)? I have \( n = 32 \) scores
Compute the test statistics.
\[ t = r \frac{n - 2}{\sqrt{1 - r^2}} = \left( 0.288 \right) \frac{30}{0.917} = 1.647 \]
use the t Distribution calculator with df=n - 2 = 30
\[ p = 0.11 \]
Interpret the results: \( p = 0.11 > 0.05 = \alpha \), do not reject \( H_0 \)

STATLAB?

For Exam 1 and STATLAB, \( r = 0.416 \). I have \( n = 32 \) scores.
Compute the test statistics.
\[ t = r \frac{n - 2}{\sqrt{1 - r^2}} = \left( 0.416 \right) \frac{30}{0.827} = 2.50 \]
use the t Distribution calculator with df=n - 2 = 30
\[ p = 0.0181 \]
Interpret the results: \( p = 0.0181 < 0.05 = \alpha \), 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.

CONCLUSIONS

- hypothesis testing
- correlation coefficient
- Fisher $z$ transform
- testing significance of correlation

NEXT TIME

- Confidence intervals with correlations
- hypothesis testing of proportions
- Can you read my mind? Part II.