

Hypothesis Testing for One Population Proportion

Requirements:

Random sample

$n \cdot p \geq 10$; $n \cdot q \geq 10$

n is sample size

Traditional Method (Non-Bayesian Testing):

Step 1: Define Claim and Opposite, H_0 (contains equal sign) and H_1

Step 2: Define significance level (α)

Step 3: Calculate Z-test statistic

$$Z_t = \frac{\hat{p} - p}{\sqrt{(p \cdot q) / n}}$$

\hat{p} - sample proportion of success

p and q are hypothetical values (success and failure)

$$q = 1 - p$$

Step 4*: Draw a picture: according to H_1 it is left-tail, right-tail or two-tail

On the picture put Z-critical value with corresponding alpha from Z-table

Step 5: Interpret results:

If Z-test statistic is in Rejection Region \Rightarrow Reject H_0 and accept H_1

If the Z-test statistic is in the Fail to Rejection Region \Rightarrow We know nothing! There is not enough evidence to accept H_1

P-value method (Bayesian Testing):

Step 1: Define Claim and Opposite, H_0 (contains equal sign) and H_1

Step 2: Define significance level (α)

Step 3: Calculate Z-test statistic

$$Z_t = \frac{\hat{p} - p}{\sqrt{(p \cdot q) / n}}$$

\hat{p} - sample proportion of success

p and q are hypothetical values (success and failure)

Step 4*: Draw a picture: according to H_1 it is left-tail, right-tail or two-tail

On the picture put Z-test statistic

P-value is : $1 - \text{Area (Fail to Rejection Region)}$; P-value is the area in the Tail;

If we have two tails we need to multiply by 2

Step 5: Interpret results:

If P-value $\leq \alpha \Rightarrow$ Reject H_0 and accept H_1

If P-value $> \alpha \Rightarrow$ Fail to Reject H_0 ; We know nothing! There is not enough evidence to accept H_1

Example in Python

For illustration, imagine that the rate of lifetime smoking in another country was known to be 40%, and we wished to assess whether the rate of lifetime smoking in the US were different from 40%.

```
condition = da["SMQ020"] == 'Yes'
```

```
da_smokers = da[condition]
```

```

n_smokers = len(da_smokers)
n_smokers

p_null = 0.40
p_hat = n_smokers / len(da["SMQ020"])
n = len(da["SMQ020"])
print(n, p_hat)
sm.stats.proportions_ztest(phat * n, n, pnull, alternative='larger')

```

Another way:

```

Z_test_statistic = (p_hat - p_null) / np.sqrt((p_null * (1 - p_null)) / n)
pvalue = 2 * dist.norm.cdf(-np.abs(Z_test_statistic))
print(Z_test_statistic, pvalue) → p1 - p2 ≠ 0

```

Hypothesis Testing for Two Population Proportions: Traditional Method (Non-Bayesian Testing)

Step 1: Define Claim and Opposite, H_0 and H_1

Step 2: Define significance level (α)

Step 3: Calculate Z-test statistic:

$$Z_t = (p1_hat - p2_hat) - (p1 - p2) / \sqrt{(p_hat_total * (1 - p_hat_total) * (1/n1 + 1/n2))}$$

$p1 - p2$ is hypothetical difference
 p_hat_total is $(X1 + X2) / (n1 + n2)$

Step 4*: Draw a picture: according to H_1 it is left-tail, right-tail or two-tail

On the picture put Z-critical value with corresponding alpha from Z-table

Step 5: Interpret results:

If Z-test statistic is in Rejection Region => Reject H_0 and accept H_1

If the Z-test statistic is in the Fail to Rejection Region => We know nothing! There is not enough evidence to accept H_1

Example in Python

if H_A : $p1 > p2$ --> $p1 - p2 > 0$, P-value = $1 - \text{dist.norm.cdf}(Z_test_stats)$

if H_A : $p1 < p2$ --> $p1 - p2 < 0$, P-value = $\text{dist.norm.cdf}(Z_test_stats)$

If H_a : $p1 - p2 \neq 0$, P-value = $2 * \text{dist.norm.cdf}(-\text{np.abs}(Z_test_statistic))$

```

Z_test_stats = best_est / np.sqrt(p_hat_total * (1 - p_hat_total) * (1/n_male_total +
1/n_female_total))

```

```

pvalue = 2*dist.norm.cdf(-np.abs(Z_test_stats)) # Multiplied by two indicates a two tailed
testing. p1 - p2 ≠ 0

```

```

print("Computed P-value is", pvalue)

```

Hypothesis Testing for One Population Mean. Population standard deviation is known

Requirements:

Random sample

Population standard deviation is known

$n > 30$ or Population is normally distributed

n is sample size

(If $n < 30$ we use T distribution)

Traditional Method (Non-Bayesian Testing):

Step 1: Define Claim and Opposite, H_0 (contains equal sign) and H_1

Step 2: Define significance level (α)

Step 3: Calculate Z-test statistic:

$$Z_t = \bar{X} - \mu / \sigma / \sqrt{n}$$

\bar{X} - sample mean

μ - population hypothetical mean

σ - population standard deviation

n - sample size

Step 4*: Draw a picture: according to H_1 it is left-tail, right-tail or two-tail

On the picture put Z-critical value with corresponding alpha from Z-table

Step 5: Interpret results:

If Z-test statistic is in Rejection Region \Rightarrow Reject H_0 and accept H_1

If the Z-test statistic is in the Fail to Rejection Region \Rightarrow We know nothing! There is not enough evidence to accept H_1

P-value method (Bayesian Testing):

Step 1: Define Claim and Opposite, H_0 (contains equal sign) and H_1

Step 2: Define significance level (α)

Step 3: Calculate Z-test statistic

$$Z_t = \bar{X} - \mu_0 / \sigma / \sqrt{n}$$

\bar{X} - sample mean

μ_0 - population hypothetical mean

σ - population standard deviation

n - sample size

Step 4*: Draw a picture: according to H_1 it is left-tail, right-tail or two-tail

On the picture put Z-test statistic

P-value is : $1 - \text{Area (Fail to Rejection Region)}$; P-value is the area in the Tail;

If we have two tails we need to multiply by 2

Step 5: Interpret results:

If P-value $\leq \alpha \Rightarrow$ Reject H_0 and accept H_1

If P-value $> \alpha \Rightarrow$ Fail to Reject H_0 ; We know nothing! There is not enough evidence to accept H_1

Example in Python:

We compare systolic blood pressure to the fixed value 120 (which is the lower threshold for "pre-hypertension")

```
sm.stats.ztest(dx.BPXS1, value=120)
```

or

```
from statsmodels.stats.weightstats import ztest
```

```
ztest_Score, p_value = ztest(data, value = null_mean, alternative='larger')
```

Hypothesis Testing for Two Population Means.

Populations are independent and standard deviations σ_1 , σ_2 are known

Requirements:

Random sample

Population standard deviation is known

$n > 30$ or Population is normally distributed

n is sample size

(If $n < 30$ we use T distribution)

Traditional Method (Non-Bayesian Testing):

Step 1: Define Claim and Opposite, H_0 and H_1

Step 2: Define significance level (α)

Step 3: Calculate Z-test statistic:

$$Z_t = \frac{(X1_{\text{bar}} - X2_{\text{bar}}) - D_0}{\sqrt{((\sigma_1^2 / n_1) + (\sigma_2^2 / n_2))}}$$

D_0 - Hypothetical difference

Step 4*: Draw a picture: according to H_1 it is left-tail, right-tail or two-tail

On the picture put Z-critical value with corresponding α from Z-table

Step 5: Interpret results:

If Z-test statistic is in Rejection Region \Rightarrow Reject H_0 and accept H_1

If the Z-test statistic is in the Fail to Rejection Region \Rightarrow We know nothing! There is not enough evidence to accept H_1

Example in Python:

```
if HA: X1 > X2 --> X1 - X2 > 0, P value = 1 - stats.t.cdf(t_test_stat, degrees_of_freedom)
```

```
if HA: X1 < X2 --> p1 - p2 < 0, P value = stats.t.cdf(t_test_stat, degrees_of_freedom)
```

```
For Equivalent --> scipy.stats.ttest_ind(a, b, equal_var=True)
```

```
bpx_female = dx.loc[dx.RIAGENDR=="Female", "BPXS1"]
```

```
bpx_male = dx.loc[dx.RIAGENDR=="Male", "BPXS1"]
```

```
print(bpx_female.mean(), bpx_male.mean())
```

```
print(sm.stats.ztest(bpx_female, bpx_male)) - for equal variances
print(sm.stats.ttest_ind(bpx_female, bpx_male)) - for equal variances
```

Hypothesis Testing for Two Population Means.

Populations are independent and standard deviations σ_1 , σ_2 are unknown

a) First approach - assumption that $(\sigma_1 \neq \sigma_2)$ is not equal to $(\sigma_2 \neq \sigma_1)$

Traditional Method (Non-Bayesian Testing)

Step 1: Define Claim and Opposite, H_0 and H_1

Step 2: Define significance level (α)

Step 3: Calculate T-test statistic:

$$T_t = (X1_bar - X2_bar) - D_0 / \sqrt{((s1^2 / n1) + (s2^2 / n2))}$$

D_0 - Hypothetical difference

$s1$ and $s2$ - sample standard deviation

Step 4*: Draw a picture: according to H_1 it is left-tail, right-tail or two-tail

On the picture put T-critical value with corresponding α from T-table and DF

Degree of Freedom = $\min(n1 - 1; n2 - 1)$; take the minimum of these two

Step 5: Interpret results:

If T-test statistic is in Rejection Region \Rightarrow Reject H_0 and accept H_1

If the T-test statistic is in the Fail to Rejection Region \Rightarrow We know nothing! There is not enough evidence to accept H_1

Example in Python

```
bmi_non_degree = da.loc[(da.RIDAGEYR >= 30) & (da.RIDAGEYR <= 40) & (da.RIAGENDR == "Female")
                        & (da.DMDEDUC2 == 0), "BMXBMI"].dropna()
bmi_degree = da.loc[(da.RIDAGEYR >= 30) & (da.RIDAGEYR <= 40) & (da.RIAGENDR == "Female")
                  & (da.DMDEDUC2 == 1), "BMXBMI"].dropna()
bmi_non_degree = sm.stats.DescrStatsW(bmi_non_degree)
bmi_degree = sm.stats.DescrStatsW(bmi_degree)
print("unequal: ", sm.stats.CompareMeans(bmi_non_degree, bmi_degree).ztest_ind(usevar='unequal',
                                                                              alternative='larger'))
```

b) Second approach - assumption that $(\sigma_1 = \sigma_2)$ is equal to $(\sigma_2 = \sigma_1)$

Traditional Method (Non-Bayesian Testing)

Step 1: Define Claim and Opposite, H_0 and H_1

Step 2: Define significance level (α)

Step 3: Calculate T-test statistic:

$$T_t = (X1_bar - X2_bar) - D_0 / \sqrt{np.sqrt((((n1 - 1) * (s1^2 / n1)) + ((n2 - 1) * (s2^2 / n2))) / (n1 + n2 - 2)) * np.sqrt((1 / n1) + (1 / n2))}$$

D_0 - Hypothetical difference

$s1$ and $s2$ - sample standard deviation

Step 4*: Draw a picture: according to H_1 it is left-tail, right-tail or two-tail

On the picture put T-critical value with corresponding α from T-table and DF

Degree of Freedom = $n1 + n2 - 2$

Step 5: Interpret results:

If T-test statistic is in Rejection Region => Reject H_0 and accept H_1

If the T-test statistic is in the Fail to Rejection Region => We know nothing! There is not enough evidence to accept H_1

Example in Python

```
bmi_non_degree = da.loc[(da.RIDAGEYR >= 30) & (da.RIDAGEYR <= 40) & (da.RIAGENDR == "Female")
                        & (da.DMDEDUC2 == 0), "BMXBMI"].dropna()
bmi_degree = da.loc[(da.RIDAGEYR >= 30) & (da.RIDAGEYR <= 40) & (da.RIAGENDR == "Female")
                  & (da.DMDEDUC2 == 1), "BMXBMI"].dropna()
bmi_non_degree = sm.stats.DescrStatsW(bmi_non_degree)
bmi_degree = sm.stats.DescrStatsW(bmi_degree)
print("pooled: ", sm.stats.CompareMeans(bmi_non_degree, bmi_degree).ztest_ind(usevar='pooled'))
```

Hypothesis Testing with Two Means.

Paired data

Traditional Method (Non-Bayesian Testing):

Step 1: Define Claim and Opposite, H_0 and H_1

Step 2: Define significance level (alpha)

Step 3: Calculate T-test statistic:

$$T_t = (X1_bar - X2_bar) - D_0 / s_d / \sqrt{n}$$

D_0 - Hypothetical difference

s_d - standard deviation difference; $S_1 - S_2$ or $\sigma_1 - \sigma_2$ (and then we have Z-test statistic)

Step 4*: Draw a picture: according to H_1 it is left-tail, right-tail or two-tail

On the picture put T-critical value with corresponding alpha from T-table and DF

Degree of Freedom = $n_1 + n_2 - 2$

Step 5: Interpret results:

If T-test statistic is in Rejection Region => Reject H_0 and accept H_1

If the T-test statistic is in the Fail to Rejection Region => We know nothing! There is not enough evidence to accept H_1

Example in Python

```
dx = da[["BPXSY1", "BPXSY2"]].dropna()
db = dx.BPXSY1 - dx.BPXSY2
print(db.mean())
sm.stats.ztest(db)
or
import scipy.stats as stats
stats.ttest_rel(pre, post)
```

Hypothesis Testing for One Population Mean. Population Standard Deviation is unknown

Requirements:

Random sample

Population standard deviation is unknown

$n > 30$ or Population is normally distributed

n is sample size

(If $n < 30$ we use T distribution too)

Traditional Method (Non-Bayesian Testing):

Step 1: Define Claim and Opposite, H_0 (contains equal sign) and H_1

Step 2: Define significance level (α)

Step 3: Calculate T-test statistic:

$$T_t = \bar{X} - \mu / s / \sqrt{n}$$

\bar{X} - sample mean

μ - population hypothetical mean

s - sample standard deviation

n - sample size

Step 4*: Draw a picture: according to H_1 it is left-tail, right-tail or two-tail

On the picture put T-critical value with corresponding α from T-table and DF

Step 5: Interpret results:

If T-test statistic is in Rejection Region \Rightarrow Reject H_0 and accept H_1

If the T-test statistic is in the Fail to Rejection Region \Rightarrow We know nothing! There is not enough evidence to accept H_1

Example in Python

We compare systolic blood pressure to the fixed value 120 (which is the lower threshold for "pre-hypertension")

```
sm.stats.ztest(dx.BPXS1, value=120)
```

Hypothesis Testing for Variance and Standard Deviation for One Population Chi-squared distribution

Traditional Method:

Step 1: Define Claim and Opposite, H_0 (contains equal sign) and H_1

Step 2: Define significance level (α)

Step 3: Calculate Test statistic:

$$\chi^2 = (n - 1) (s^2 / \sigma^2)$$

Step 4*: Draw a picture: according to H_1 it is left-tail, right-tail or two-tail; Keep in mind that distribution starts from Zero and it's only Right-Skewed

On the picture put Chi-squared critical value with corresponding alpha and DF

Example in Python

$n = 12$, 95% Confidence Level, Degrees of Freedom = 11 and test for Equality

`right_critical = scipy.stats.chi2.ppf(1-.025, df = 11)` is $XR2 = 21.92$

`left_critical = scipy.stats.chi2.ppf(.025, df=11)` is $XL2 = 3,816$

Step 5: Interpret results:

If the Test statistic is in Rejection Region \Rightarrow Reject H_0 and accept H_1

If the Test statistic is in the Fail to Rejection Region \Rightarrow We know nothing! There is not enough evidence to accept H_1

Hypothesis Testing for Comparing Two Variances - Two independent Samples and populations are normally distributed

F- distribution

Step 1: Define Claim and Opposite, $H_0 (F = 1)$ and $H_1 (\sigma^2_1 > \sigma^2_2)$

Step 2: Define significance level (alpha)

Step 3: Calculate Test statistic:

$F = \text{variance}_{1(\text{larger})} / \text{variance}_{2(\text{smaller})}$

Step 4*: Draw a picture: it is always upper-tailed

On the picture put F-critical value which is based on DF of these two samples and alpha

`import scipy.stats`

`print(scipy.stats.f.ppf(q=1-.005, dfn=4, dfd=6))`

alpha = 0.05 for 95% Confidence level

DF numerator = $n_1 - 1$

DF denominator = $n_2 - 1$

Step 5: Interpret results:

If Test statistic is in Rejection Region \Rightarrow Reject $H_0 (F = 1)$ and accept H_1

If the Test statistic is in the Fail to Rejection Region \Rightarrow We know nothing! There is not enough evidence to accept H_1

Hypothesis Testing: Chi-Square Test

It helps us to understand the relationship between two categorical variables:

grade level, sex, age group, year. Chi-Square test involve the frequency of events; the count; Expected Vs Observed categorical distribution

Example - determine if this die is a fair or not with 95% certainty; 600 trials for the next 6 days;

Example - a school principal expected that students will be absent equally during the 5-day school week;

Step 1: Construct two tables: Observed vs Expected value

Step 2: Define Claim and Opposite, H_0 (contains equal sign) and H_1

Step 3: Define significance level (alpha)

Step 4: Calculate Test statistic:

$$\text{Chi-Squared} = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$$

Step 5: Draw a picture:

Keep in mind that distribution starts from Zero and it's only Right-Skewed

On the picture put Chi-Squared critical value which is based on $DF(n - 1)$ and alpha, which is calculated with calculator (or with Python)

Step 6: Interpret results:

If Test statistic is in Rejection Region \Rightarrow Reject H_0 and accept H_1

if the Test statistic is in the Fail to Rejection Region \Rightarrow We know nothing! There is not enough evidence to accept H_1

Test of Independence Using Chi-Square Distribution

(watch the video again in your playlist)

Example: Is the average number of studying hours depend on the type of student;

Step 1: Construct two tables: Observed vs Expected value ($E.V = \text{Row}_{\text{total}} * \text{Col}_{\text{total}} / N$)

Step 2: Define Claim and Opposite, H_0 (contains equal sign) and H_1

Step 3: Define significance level (alpha)

Step 4: Calculate Test statistic:

$$\text{Chi-Squared} = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$$

Step 5: Draw a picture:

Keep in mind that distribution starts from Zero and it's only Right-Skewed

On the picture put Chi-Squared critical value which is based on $DF(n - 1)$ and alpha, which is calculated with calculator (or with Python)

Step 6: Interpret results:

If Test statistic is in Rejection Region \Rightarrow Reject H_0 and accept H_1

if the Test statistic is in the Fail to Rejection Region \Rightarrow We know nothing! There is not enough evidence to accept H_1