Week 6:
Statistical tests for one variable

Phase 3

Key concepts

- Sometimes we want to test **our estimated population** (e.g. the mean) against a particular value (our null hypothesis).
- This is usually called **one sample** statistical testing.
- Different **data types** and **ASSUMPTIONS** → **a particular** statistical test.
- Here (and in all phase III), we are going to concentrate on **parametric** tests, which assumes an **underlying distribution** to compute the p-values.

Key concepts

- Sometimes we want to test **our estimated population** (e.g. the mean) against a particular value (our null hypothesis).
- This is usually called **one sample** statistical testing.
- Different **data types** and **ASSUMPTIONS** → **a particular** statistical test.
- Here (and in all phase III), we are going to concentrate on **parametric** tests, which assumes an **underlying distribution** to compute the p-values.

Your dataset

Inferences based on the values of one variable, which represents an observed property of the population.

gender	education	age	ACT	SATV	SATQ
2	3	19	24	500	500
2	3	23	35	600	500
2	3	20	21	480	470
1	4	27	26	550	520
1	2	33	31	600	550
1	5	26	28	640	640
2	5	30	36	610	500
1	3	19	22	520	560
2	4	23	22	400	600
2	5	40	35	730	800

Your dataset

Inferences based on the values of one variable, which represents an observed property of the population.

e.g. is the average ACT score of the population greater than 25?

gender	education	age	ACT	SATV	SATQ
2	3	19	24	500	500
2	3	23	35	600	500
2	3	20	21	480	470
1	4	27	26	550	520
1	2	33	31	600	550
1	5	26	28	640	640
2	5	30	36	610	500
1	3	19	22	520	560
2	4	23	22	400	600
2	5	40	35	730	800

Your dataset

Inferences based on the values of one variable, which represents an observed property of the population.

e.g. does a population that takes ACT exams differ in the frequency of education levels?

gender	education	age	ACT	SATV	SATQ
2	3	19	24	500	500
2	3	23	35	600	500
2	3	20	21	480	470
1	4	27	26	550	520
1	2	33	31	600	550
1	5	26	28	640	640
2	5	30	36	610	500
1	3	19	22	520	560
2	4	23	22	400	600
2	5	40	35	730	800

Tests for a continuous variable

- Here you'll probably want to test the estimation of the mean μ in a certain property/aspect of the population with respect to a hypothesized value (μ_0).

$$\mathbf{H_0}: \mu=\mu_0$$

$$\mathbf{H_A}: \mu>\mu_0 \text{ or } \mu < \mu_0 \text{ (one-sided)}$$

$$\mu \neq \mu_0 \text{ (two-sided)}$$

- <u>Example</u>: "is the insula active ($\mu \neq 0$) during stressor-evoked tasks?"

- Depending on the scenario (+assumptions), you may run either a one sample **z-test** or **t-test.**

One sample z-test

Assumptions:

1. Independence: Observations in your sample are not correlated with each other (e.g. in *cross-sectional* studies).

2. Known standard deviation. We know the standard deviation of the population.

3. Normality. Our variable follows a gaussian distribution or sample sizes are big enough (thanks, Central Limit Theorem!)

One sample z-test

- Null Hypothesis $\mathbf{H_0}$: $\mu=\mu_0$

- Test statistic: $z = \dfrac{< X > - \mu_0}{\sigma/\sqrt{N}}$

- Alternative Hypothesis $\mathbf{H_A}$

  $\mu > \mu_0$ (one-sided right tail)

  $\mu < \mu_0$ (one-sided left tail)

  $\mu \neq \mu_0$ (two-sided)

One sample z-test

- Null Hypothesis $\mathbf{H_0}$: $\mu=\mu_0$

- Test statistic: $z = \dfrac{< X > - \mu_0}{\sigma/\sqrt{N}}$ ~ Normal(0, 1)

- Alternative Hypothesis $\mathbf{H_A}$

  $\mu > \mu_0$ (one-sided right tail)

  $\mu < \mu_0$ (one-sided left tail)

  $\mu \neq \mu_0$ (two-sided)

One sample z-test

- Null Hypothesis $\mathbf{H_0}$: $\mu=\mu_0$

- Test statistic: $z = \dfrac{<X> - \mu_0}{\sigma/\sqrt{N}}$ ~ Normal(0, 1)

- Alternative Hypothesis $\mathbf{H_A}$ <u>Rejection region for α</u>

  $\mu > \mu_0$ (one-sided right tail) $\longrightarrow$ $z \geq z_\alpha$

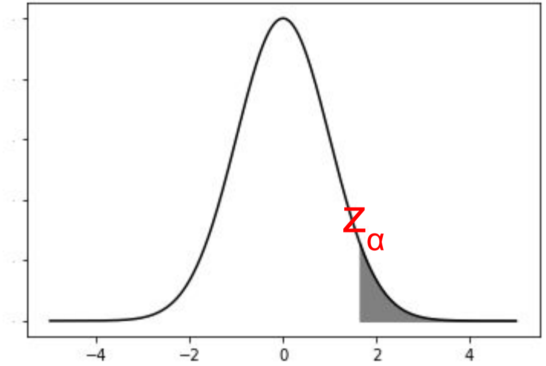
  $\mu < \mu_0$ (one-sided left tail) $\longrightarrow$ $z \leq z_\alpha$

  $\mu \neq \mu_0$ (two-sided) $\longrightarrow$ $|z| \geq |z_{\alpha/2}|$

One sample z-test

- Null Hypothesis $\mathbf{H_0}$: $\mu=\mu_0$

- Test statistic: $z = \dfrac{< X > - \mu_0}{\sigma/\sqrt{N}}$



- Alternative Hypothesis $\mathbf{H_A}$

<u>Rejection region for α</u>

$\mu > \mu_0$ (one-sided right tail) $\longrightarrow$ $z \geq z_\alpha$

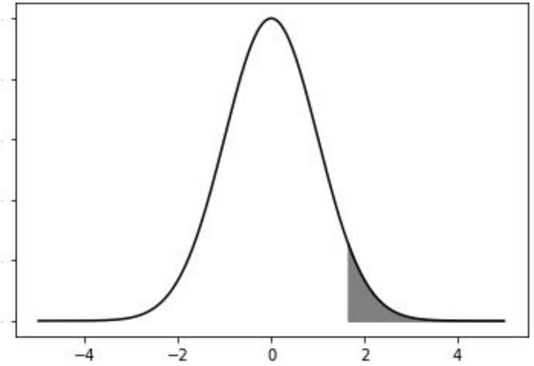
qnorm(α, lower.tail =FALSE)

$\mu < \mu_0$ (one-sided left tail) $\longrightarrow$ $z \leq z_\alpha$

$\mu \neq \mu_0$ (two-sided) $\longrightarrow$ $|z| \geq |z_{\alpha/2}|$

One sample z-test

- Null Hypothesis $\mathbf{H_0}$: $\mu=\mu_0$



- Test statistic: $z = \dfrac{<X> - \mu_0}{\sigma/\sqrt{N}}$

- Alternative Hypothesis $\mathbf{H_A}$ <u>Rejection region for α</u> (in P-VALUES)

$\mu > \mu_0$ (one-sided right tail) $\longrightarrow$ $P(Z \geq z \mid H_0) \leq \alpha$

$\mu < \mu_0$ (one-sided left tail) $\longrightarrow$ $z \leq z_\alpha$

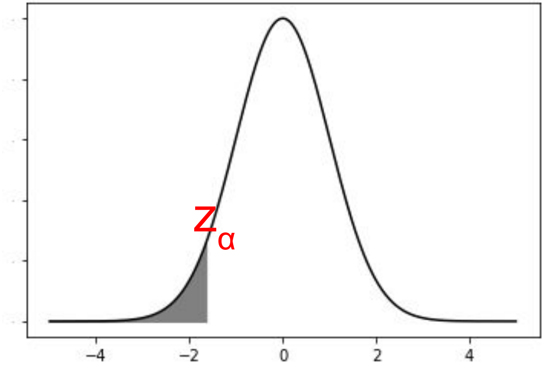
$\mu \neq \mu_0$ (two-sided) $\longrightarrow$ $|z| \geq |z_{\alpha/2}|$

pnorm(z, lower.tail =FALSE)

One sample z-test

- Null Hypothesis $\mathbf{H_0}$: $\mu=\mu_0$

- Test statistic: $$z = \frac{< X > - \mu_0}{\sigma/\sqrt{N}}$$



- Alternative Hypothesis $\mathbf{H_A}$ — <u>Rejection region for α</u>

$\mu > \mu_0$ (one-sided right tail) ⟶ $z \geq z_\alpha$

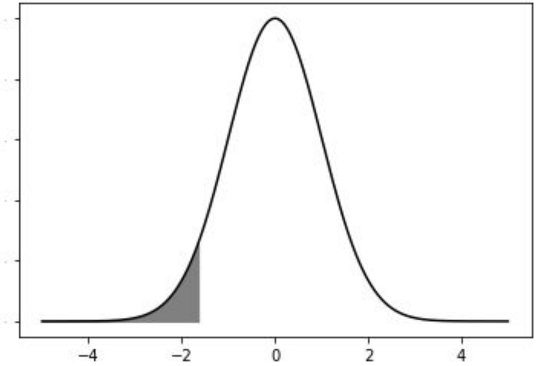
$\mu < \mu_0$ (one-sided left tail) ⟶ $z \leq z_\alpha$

$\mu \neq \mu_0$ (two-sided) ⟶ $|z| \geq |z_{\alpha/2}|$

qnorm(α, lower.tail =TRUE)

One sample z-test

- Null Hypothesis $\mathbf{H_0}$: $\mu=\mu_0$



- Test statistic: $z = \dfrac{<X> - \mu_0}{\sigma/\sqrt{N}}$

- Alternative Hypothesis $\mathbf{H_A}$ — <u>Rejection region for α</u> (in P-VALUES)

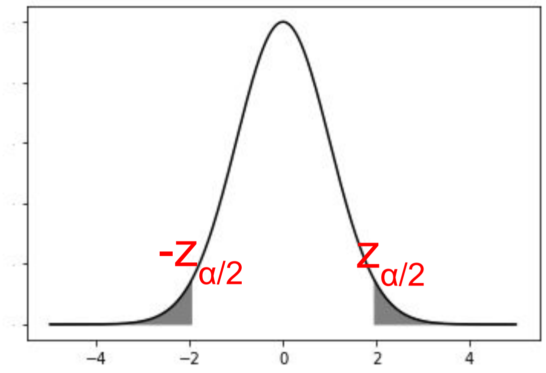
$\mu > \mu_0$ (one-sided right tail) $\longrightarrow$ $z \geq z_\alpha$

$\mu < \mu_0$ (one-sided left tail) $\longrightarrow$ $P(Z \leq z \mid H_0) \leq \alpha$ — *pnorm*(z, lower.tail =TRUE)

$\mu \neq \mu_0$ (two-sided) $\longrightarrow$ $|z| \geq |z_{\alpha/2}|$

One sample z-test

- Null Hypothesis $\mathbf{H_0}$: $\mu=\mu_0$



- Test statistic: $z = \dfrac{<X> - \mu_0}{\sigma/\sqrt{N}}$

- Alternative Hypothesis $\mathbf{H_A}$ — <u>Rejection region for α</u>

$\mu > \mu_0$ (one-sided right tail) $\longrightarrow$ $z \geq z_\alpha$ — *qnorm*(α/2, lower.tail =TRUE)

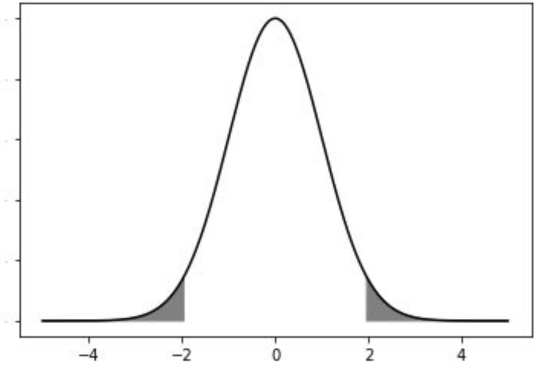
$\mu < \mu_0$ (one-sided left tail) $\longrightarrow$ $z \leq z_\alpha$ — *qnorm*(α/2, lower.tail =FALSE)

$\mu \neq \mu_0$ (two-sided) $\longrightarrow$ $|z| \geq |z_{\alpha/2}|$

One sample z-test

- Null Hypothesis $\mathbf{H_0}$: $\mu=\mu_0$

- Test statistic: $$z = \frac{<X> - \mu_0}{\sigma/\sqrt{N}}$$



- Alternative Hypothesis $\mathbf{H_A}$ Rejection region for α (in P-VALUES)

$\mu > \mu_0$ (one-sided right tail) $\longrightarrow$ $z \geq z_\alpha$

$\mu < \mu_0$ (one-sided left tail) $\longrightarrow$ $z \leq z_\alpha$

$\mu \neq \mu_0$ (two-sided) $\longrightarrow$ $P(Z \geq |z| \mid H_0) \leq \alpha$

2**pnorm*(|z|, lower.tail =FALSE)

Practice example

The melting point of each of 16 samples of a certain brand of hydrogenated vegetable oil was determined, resulting in $\langle X \rangle = 94.32$. Assume that the distribution of melting point is normal with $\sigma = 1.20$. Can we reject that the population melting point of hydrogenated vegetable oil, μ, is equal to 95 (i.e. our null hypothesis) at a $\alpha = 0.01$ significance level (Hint: $|z_{0.01/2}| \sim 2.57$)?

One sample t-test

Assumptions:

1. Independence: Observations in your sample are not correlated with each other (e.g. in *cross-sectional* studies).

2. Unknown standard deviation. We don't know the standard deviation of the population and thus, we use the estimated one, $\hat{\sigma}$, from our sample.

3. Normality. Our variable follows a gaussian distribution or sample sizes are big enough (thanks, Central Limit Theorem!)

One sample t-test

- Null Hypothesis $\mathbf{H_0}$: $\mu=\mu_0$

- Test statistic: $t = \dfrac{< X > - \mu_0}{\hat{\sigma}/\sqrt{N}}$

- Alternative Hypothesis $\mathbf{H_A}$

  $\mu > \mu_0$ (one-sided right tail)

  $\mu < \mu_0$ (one-sided left tail)

  $\mu \neq \mu_0$ (two-sided)

One sample t-test

- Null Hypothesis $\mathbf{H_0}$: $\mu=\mu_0$

- Test statistic: $t = \dfrac{< X > -\mu_0}{\hat{\sigma}/\sqrt{N}}$ ~ Student's t (df=N-1)

- Alternative Hypothesis $\mathbf{H_A}$

  $\mu > \mu_0$ (one-sided right tail)

  $\mu < \mu_0$ (one-sided left tail)

  $\mu \neq \mu_0$ (two-sided)

One sample t-test

- Null Hypothesis $\mathbf{H_0}$: $\mu=\mu_0$
- Test statistic: $t = \dfrac{<X> - \mu_0}{\hat{\sigma}/\sqrt{N}}$ ~ Student's t (df=N-1)
- Alternative Hypothesis $\mathbf{H_A}$ — <u>Rejection region for α</u>

Alternative Hypothesis $\mathbf{H_A}$	Rejection region for α
$\mu > \mu_0$ (one-sided right tail)	$t \geq t_\alpha$
$\mu < \mu_0$ (one-sided left tail)	$t \leq t_\alpha$
$\mu \neq \mu_0$ (two-sided)	$\|t\| \geq \|t_{\alpha/2}\|$

One sample t-test

- Null Hypothesis $\mathbf{H_0}$: $\mu=\mu_0$

- Test statistic: $t = \dfrac{< X > - \mu_0}{\hat{\sigma}/\sqrt{N}}$

t.test(X, mu = μ_0, alternative=”greater”)

- Alternative Hypothesis $\mathbf{H_A}$

Rejection region for α (in P-VALUES)

Alternative Hypothesis $\mathbf{H_A}$	Rejection region for α (in P-VALUES)
$\mu > \mu_0$ (one-sided right tail)	$P(T \geq t \mid H_0) \leq \alpha$
$\mu < \mu_0$ (one-sided left tail)	$t \leq t_\alpha$
$\mu \neq \mu_0$ (two-sided)	$\lvert t \rvert \geq \lvert t_{\alpha/2} \rvert$

One sample t-test

- Null Hypothesis $\mathbf{H_0}$: $\mu=\mu_0$

- Test statistic: $t = \frac{<X> - \mu_0}{\hat{\sigma}/\sqrt{N}}$

t.test(X, mu = μ_0, alternative="less")

- Alternative Hypothesis $\mathbf{H_A}$ — Rejection region for α (in P-VALUES)

$\mu > \mu_0$ (one-sided right tail)	$t \geq t_\alpha$
$\mu < \mu_0$ (one-sided left tail)	$P(T \leq t \mid H_0) \leq \alpha$
$\mu \neq \mu_0$ (two-sided)	$\lvert t\rvert \geq \lvert t_{\alpha/2}\rvert$

One sample t-test

- Null Hypothesis $\mathbf{H_0}$: $\mu=\mu_0$
- Test statistic: $t = \dfrac{<X> - \mu_0}{\hat{\sigma}/\sqrt{N}}$

$\boxed{\textit{t.test}(X,\ mu = \mu_0,\ alternative=\text{"two.sided"})}$

- Alternative Hypothesis $\mathbf{H_A}$ — <u>Rejection region for α</u> (in P-VALUES)

Alternative Hypothesis $\mathbf{H_A}$	Rejection region for α (in P-VALUES)
$\mu > \mu_0$ (one-sided right tail)	$t \geq t_\alpha$
$\mu < \mu_0$ (one-sided left tail)	$t \leq t_\alpha$
$\mu \neq \mu_0$ (two-sided)	$P(T \geq \|t\| \mid H_0) \leq \alpha$

Tests for a categorical variable

- Here you'll probably want to test how likely the population proportions of *i* categories are with respect to some hypothesized values (p_i).

$$\mathbf{H_0}: p_i = p_{i0}$$
$$\mathbf{H_A}: p_i \neq p_{i0}$$

- Example: You roll a dice *N* times. If the dice is fair, you should get more or less a similar proportion across all faces (1, 2, 3, 4, 5, 6).

- Depending on the scenario (+assumptions), you may run either a χ^2 **test** or a **binomial test.**

χ^2 test

- It aims to test differences in observed proportions p_i with respect to some hypothesized p_{i0} across k categories.
- For this case (one variable), also known as a Pearson's χ^2 test or **goodness-of-fit test**.

χ^2 test

- It aims to test differences in observed proportions p_i with respect to some hypothesized p_{i0} across k categories.
- For this case (one variable), also known as a Pearson's χ^2 test or **goodness-of-fit test**.

Assumptions:

1. Independence: Observations in your sample are not correlated with each other (e.g. in *cross-sectional* studies).
2. Sufficiently large expected frequencies. Rule of thumb: at least 5 expected observations in each category.

χ^2 test

- Null Hypothesis $\mathbf{H_0}$: $p_i = p_{i0}$

- Test statistic: $$\chi^2 = \sum_{i=1}^{k} \left[\frac{(n_i - E_i)^2}{E_i} \right]$$

- Alternative Hypothesis $\mathbf{H_A}$:

  $p_i \neq p_{i0}$ (In any category)

χ^2 test

- Null Hypothesis **H_0**: $p_i = p_{i0}$

- Test statistic: $$\chi^2 = \sum_{i=1}^{k} \left[\frac{(n_i - E_i)^2}{E_i} \right]$$

  n_i = number of observations in category *i*
  $E_i = N \times p_i$ = *expected number under H_0*

- Alternative Hypothesis **H_A**:

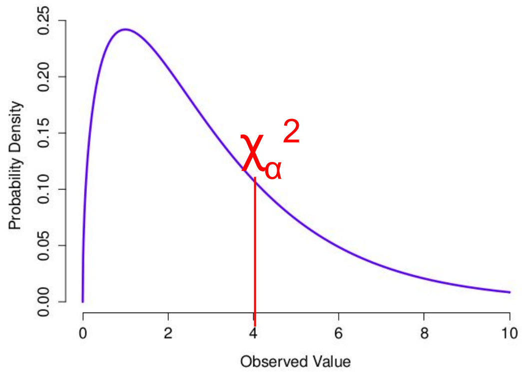
  $p_i \neq p_{i0}$ (In any category)

χ^2 test

- Null Hypothesis $\mathbf{H_0}$: $p_i = p_{i0}$

- Test statistic: $$\chi^2 = \sum_{i=1}^{k} \left[\frac{(n_i - E_i)^2}{E_i} \right]$$

n_i = number of observations in category i
$E_i = N \times p_i$ = *expected number under* H_0



qchisq(α/2, df=*k-1*, lower.tail =FALSE)

- Alternative Hypothesis $\mathbf{H_A}$:

$p_i \neq p_{i0}$ (In any category)

<u>Rejection region for α</u>

$\chi^2 \geq \chi^2_{\alpha,k-1}$

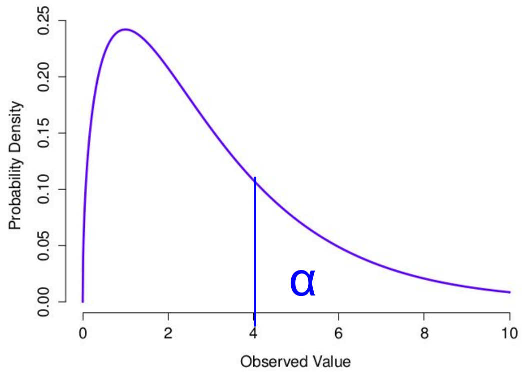
χ^2 test

- Null Hypothesis $\mathbf{H_0}$: $p_i = p_{i0}$

- Test statistic: $$\chi^2 = \sum_{i=1}^{k} \left[\frac{(n_i - E_i)^2}{E_i} \right]$$

n_i = number of observations in category i

$E_i = N \times p_i$ = *expected number under* H_0



pchisq(χ^2, df=*k-1*, lower.tail =FALSE)

- Alternative Hypothesis $\mathbf{H_A}$:

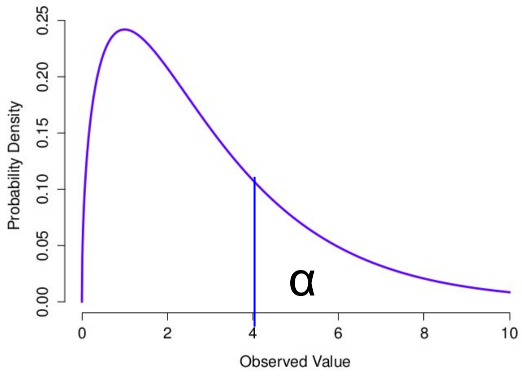
$p_i \neq p_{i0}$ (In any category)

<u>Rejection region for α</u>

$P(X \geq \chi^2) \leq \alpha$

χ^2 test

- Null Hypothesis $\mathbf{H_0}$: $p_i = p_{i0}$



- Test statistic: $$\chi^2 = \sum_{i=1}^{k} \left[\frac{(n_i - E_i)^2}{E_i} \right]$$

n_i = number of observations in category i
$E_i = N \times p_i$ = *expected number under H_0*

The χ^2 is always a one-sided test!!!

- Alternative Hypothesis $\mathbf{H_A}$:

$p_i \neq p_{i0}$ (In any category)

<u>Rejection region for α</u>

$P(X \geq \chi^2) \leq \alpha$

χ^2 test

- Null Hypothesis $\mathbf{H_0}$: $p_i = p_{i0}$

- Test statistic: $$\chi^2 = \sum_{i=1}^{k} \left[\frac{(n_i - E_i)^2}{E_i} \right]$$

$\boxed{\textit{chisq.test}(X, p = p_{i0})}$

n_i = number of observations in category *i*
$E_i = N \times p_i$ = *expected number under* H_0

- Alternative Hypothesis $\mathbf{H_A}$:

$p_i \neq p_{i0}$ (In any category)

<u>Rejection region for α</u>

$P(X \geq \chi^2) \leq \alpha$

Practice example

A random sample of 150 recent donations at a certain blood bank reveals that 82 were type A blood. At a significance level $\alpha = 0.05$, does this suggest that the actual percentage of type A donations differs from 40%, the percentage of the population having type A blood? (Hint: $\chi^2_{0.05,1} \sim 3.84$)

Binomial test

- The χ^2 test works reasonably well when there are sufficiently large observations expected in each category.
- If our variable has only two categories (e.g. a coin) we could use the **binomial test**.
- <u>Example</u>: testing the fairness of a coin (equal proportion of heads and tails, 0.5) based on my observations.
- Here probabilities are exact, so for small samples, it might be more powerful than the χ^2 test.
- It's a **two-tailed** test: we may test for greater, less, or unequal.

Binomial test

Assumptions:

1. Independence: Observations in your sample are not correlated with each other (e.g. in *cross-sectional* studies).
2. Only 2 categories. This test only works when your variable has 2 categories.

Binomial test

- Null Hypothesis $\mathbf{H_0}$: $p=p_0$
- (Exact) Probability of getting *l* outcomes: $P(X=l|\,p)=\binom{N}{l}p^l(1-p)^{N-l}$
- Alternative Hypothesis $\mathbf{H_A}$:

  $p > p_0$

  $p < p_0$

  $p \neq p_0$

Binomial test

- Null Hypothesis $\mathbf{H_0}$: $p=p_0$

- (Exact) Probability of getting *l* outcomes: $P(X = l|\ p) = \binom{N}{l} p^l (1-p)^{N-l}$

- Alternative Hypothesis $\mathbf{H_A}$:

  $p > p_0$

  $p < p_0$

  $p \neq p_0$

<u>Rejection condition for α</u>

$$\sum_{i=l}^{N} P(X = i|\ p = p_0) \leq \alpha$$

Binomial test

- Null Hypothesis $\mathbf{H_0}$: $p=p_0$

- (Exact) Probability of getting *l* outcomes: $P(X = l|\, p) = \binom{N}{l} p^l (1-p)^{N-l}$

- Alternative Hypothesis $\mathbf{H_A}$:

<u>Rejection condition for α</u>

$p > p_0$

$p < p_0$

$p \neq p_0$

$$\sum_{i=l}^{N} P(X = i|\, p = p_0) \leq \alpha$$

binom.test(l, N, p = p_0, alternative = "less")

Binomial test

- Null Hypothesis $\mathbf{H_0}$: $p=p_0$

- (Exact) Probability of getting *l* outcomes: $P(X=l|\ p) = \binom{N}{l} p^l (1-p)^{N-l}$

- Alternative Hypothesis $\mathbf{H_A}$:

<u>Rejection condition for α</u>

$p > p_0$

$p < p_0$

$p \neq p_0$

$$\sum_{i=0}^{l} P(X=i|\ p=p_0) \leq \alpha$$

Binomial test

- Null Hypothesis $\mathbf{H_0}$: $p=p_0$

- (Exact) Probability of getting *l* outcomes: $P(X = l|\ p) = \binom{N}{l} p^l (1-p)^{N-l}$

- Alternative Hypothesis $\mathbf{H_A}$:

Rejection condition for α

$p > p_0$

$p < p_0$

$p \neq p_0$

$$\sum_{i=0}^{l} P(X = i|\ p = p_0) \leq \alpha$$

binom.test(l, N, p = p_0, alternative = “greater”)

Binomial test

- Null Hypothesis $\mathbf{H_0}$: $p=p_0$

- (Exact) Probability of getting *l* outcomes: $P(X = l|\ p) = \binom{N}{l} p^l (1-p)^{N-l}$

- Alternative Hypothesis $\mathbf{H_A}$:

Rejection condition for α

$p > p_0$

$p < p_0$

$p \neq p_0$

binom.test(l, N, p = p_0, alternative = "two.sided")

Recap

- ➢ Sometimes we may want to test **one aspect** of the population with respect to a **particular value**.
- ➢ This is usually done by just applying a **statistical test to one of our variables** in the data.
- ➢ That variable will determine which **appropriate statistical test** to use.
- ➢ In the next few weeks, we will see tests that involve **two (and even more) variables** in our data.