z-test

- **Use:** Compare the data mean to an hypothesized mean.
- **Data:** \(x_1, x_2, \ldots, x_n\).
- **Assumptions:** The data are independent normal samples: \(x_i \sim N(\mu, \sigma^2)\) where \(\mu\) is unknown, but \(\sigma\) is known.
- **\(H_0\):** For a specified \(\mu_0\), \(\mu = \mu_0\).
- **\(H_A\):**
  - Two-sided: \(\mu \neq \mu_0\)
  - One-sided-greater: \(\mu > \mu_0\)
  - One-sided-less: \(\mu < \mu_0\)
- **Test statistic:**
  \[ z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \]
- **Null distribution:** \(\phi(z | H_0)\) is the pdf of \(Z \sim N(0, 1)\).
- **\(p\)-value:**
  - Two-sided:
    \[ p = P(|Z| > z | H_0) = 2 \cdot (1 - \text{pnorm}(\text{abs}(z), 0, 1)) \]
  - One-sided-greater (right-sided):
    \[ p = P(Z > z | H_0) = 1 - \text{pnorm}(z, 0, 1) \]
  - One-sided-less (left-sided):
    \[ p = P(Z < z | H_0) = \text{pnorm}(z, 0, 1) \]
- **Critical values:** \(z_\alpha\) has right-tail probability \(\alpha\)
  \[ P(z > z_\alpha | H_0) = \alpha \iff z_\alpha = \text{qnorm}(1 - \alpha, 0, 1). \]
- **Rejection regions:** let \(\alpha\) be the significance.
  - Right-sided rejection region: \([z_\alpha, \infty)\)
  - Left-sided rejection region: \((1 - \alpha, z_\alpha]\)
  - Two-sided rejection region: \((1 - \alpha, \text{qnorm}(1 - \alpha, 0, 1)) \cup [z_\alpha, \infty)\)

**Alternate test statistic**

- **Test statistic:** \(\bar{x}\)

- **Null distribution:** \(\phi(\bar{x} | H_0)\) is the pdf of \(\bar{X} \sim N(\mu_0, \sigma^2/n)\).
- **\(p\)-value:**
  - Two-sided:
    \[ p = P(|\bar{X} - \mu_0| > |\bar{x} - \mu_0| | H_0) = 2 \cdot (1 - \text{pnorm}(\text{abs}((\bar{x} - \mu_0), 0, \sigma/\sqrt{n}))) \]
  - One-sided-greater: \(\mu > \mu_0\)
    \[ p = P(\bar{X} > \bar{x}) = 1 - \text{pnorm}(\bar{x}, \mu_0, \sigma/\sqrt{n}) \]
  - One-sided-less:
    \[ p = P(\bar{X} < \bar{x}) = \text{pnorm}(\bar{x}, \mu_0, \sigma/\sqrt{n}) \]
- **Critical values:** \(x_\alpha\) has right-tail probability \(\alpha\)
  \[ P(X > x_\alpha | H_0) = \alpha \iff x_\alpha = \text{qnorm}(1 - \alpha, \mu_0, \sigma/\sqrt{n}) \]
- **Rejection regions:** let \(\alpha\) be the significance.
  - Right-sided rejection region: \([x_\alpha, \infty)\)
  - Left-sided rejection region: \((1 - \alpha, x_\alpha]\)
  - Two-sided rejection region: \((1 - \alpha, \text{qnorm}(1 - \alpha, \mu_0, \sigma/\sqrt{n})) \cup [x_\alpha, \infty)\)
One-sample $t$-test of the mean

- Use: Compare the data mean to an hypothesized mean.
- Data: $x_1, x_2, ..., x_n$.
- Assumptions: The data are independent normal samples: 
  $x_i \sim N(\mu, \sigma^2)$ where both $\mu$ and $\sigma$ are unknown.
- $H_0$: For a specified $\mu_0$, $\mu = \mu_0$
- $H_A$:
  Two-sided: $\mu \neq \mu_0$
  one-sided-greater: $\mu > \mu_0$
  one-sided-less: $\mu < \mu_0$
- Test statistic: 
  $$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}},$$
  where $s^2$ is the sample variance: 
  $$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$
- Null distribution: $\phi(t \mid H_0)$ is the pdf of $T \sim t(n-1)$.
  (Student $t$-distribution with $n-1$ degrees of freedom)
- $p$-value:
  Two-sided: $p = P(|T| > t) = 2(1 - pt(abs(t), n-1))$
  one-sided-greater: $p = P(T > t) = 1 - pt(t, n-1)$
  one-sided-less: $p = P(T < t) = pt(t, n-1)$
- Critical values: $t_\alpha$ has right-tail probability $\alpha$
  $$P(T > t_\alpha \mid H_0) = \alpha \Leftrightarrow t_\alpha = qt(1 - \alpha, n-1).$$
- Rejection regions: let $\alpha$ be the significance.
  Right-sided rejection region: $[t_\alpha, \infty)$
  Left-sided rejection region: $(-\infty, t_{1-\alpha}]$
  Two-sided rejection region: $(-\infty, t_{1-\alpha/2}] \cup [t_{\alpha/2}, \infty)$

Two-sample $t$-test for comparing means (assuming equal variance)

- Use: Compare the means from two groups.
- Data: $x_1, x_2, ..., x_n$ and $y_1, y_2, ..., y_m$.
- Assumptions: Both groups of data are independent normal samples:
  $$x_i \sim N(\mu_x, \sigma^2)$$
  $$y_j \sim N(\mu_y, \sigma^2)$$
  where both $\mu_x$ and $\mu_y$ are unknown and possibly different. The variance $\sigma$ is unknown,
  but the same for both groups.
- $H_0$: $\mu_x = \mu_y$
- $H_A$:
  Two-sided: $\mu_x \neq \mu_y$
  one-sided-greater: $\mu_x > \mu_y$
  one-sided-less: $\mu_x < \mu_y$
Summary of NHST for 18.03, Spring 2022

- Test statistic: $t = \frac{\bar{x} - \bar{y}}{s_p}$,
  where $s^2_x$ and $s^2_y$ are the sample variances and $s^2_p$ is (sometimes called) the pooled sample variance:
  
  $$s^2_p = \frac{(n - 1)s^2_x + (m - 1)s^2_y}{n + m - 2} \left(\frac{1}{n} + \frac{1}{m}\right)$$

- Null distribution: $\phi(t \mid H_0)$ is the pdf of $T \sim t(n + m - 2)$. (Student $t$-distribution with $n + m - 2$ degrees of freedom.)

- $p$-value:
  Two-sided: $p = P(|T| > t) = 2(1 - pt(abs(t), n+m-2))$
  one-sided-greater: $p = P(T > t) = 1 - pt(t, n+m-2)$
  one-sided-less: $p = P(T < t) = pt(t, n+m-2)$

- Critical values: $t_\alpha$ has right-tail probability $\alpha$

  $$P(t > t_\alpha \mid H_0) = \alpha \iff t_\alpha = qt(1 - \alpha, n + m - 2).$$

- Rejection regions: let $\alpha$ be the significance.
  Right-sided rejection region: $[t_\alpha, \infty)$
  Left-sided rejection region: $(-\infty, t_{1-\alpha}]$
  Two-sided rejection region: $(-\infty, t_{1-\alpha/2}] \cup [t_{\alpha/2}, \infty)$

Notes: 1. Unequal variances. There is a form of the $t$-test for when the variances are not assumed equal. It is sometimes called Welch’s $t$-test. In the R function t.test, there is an argument var.equal. Setting it to FALSE runs the unequal variances version of the $t$-test.

2. When the data naturally comes in pairs $(x_i, y_i)$, one uses the paired two-sample $t$-test. For example, in comparing two treatments, each patient receiving treatment 1 might be paired with a patient receiving treatment 2 who is similar in terms of stage of disease, age, sex, etc.

$^2$ test for variance

- Use: Compare the data variance to an hypothesized variance.
- Data: $x_1, x_2, \ldots, x_n$.
- Assumptions: The data are independent normal samples:

  $x_i \sim N(\mu, \sigma^2)$ where both $\mu$ and $\sigma$ are unknown.
- $H_0$: For a specified $\sigma_0$, $\sigma = \sigma_0$
- $H_A$:

  Two-sided: $\sigma \neq \sigma_0$
  one-sided-greater: $\sigma > \sigma_0$
  one-sided-less: $\sigma < \sigma_0$

- Test statistic: $X^2 = \frac{(n - 1)s^2}{\sigma^2_0}$, where $s^2$ is the sample variance:

  $$s^2 = \frac{1}{n - 1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$
• Null distribution: \( \phi(X^2 | H_0) \) is the pdf of \( \chi^2 \sim \chi^2(n-1) \).
  (Chi-square distribution with \( n - 1 \) degrees of freedom)

• \( p \)-value:  
  Because the \( \chi^2 \) distribution is not symmetric around zero the two-sided test is a little awkward to write down. The idea is to look at the \( X^2 \) statistic and see if it’s in the left or right tail of the distribution. The \( p \)-value is twice the probability in that tail.

  An easy check for which tail it’s in is: \( s^2/\sigma_0^2 > 1 \) (right tail) or \( s^2/\sigma_0^2 < 1 \) (left tail).

\[
\text{Two-sided: } p = \begin{cases} 
2 \times P(\chi^2 > X^2) & \text{if } X^2 \text{ is in the right tail} \\
2 \times P(\chi^2 < X^2) & \text{if } X^2 \text{ is in the left tail} 
\end{cases} = 2 \times \min(p\text{chisq}(X^2,n-1), 1-p\text{chisq}(X^2,n-1))
\]

\[
\text{one-sided-greater: } p = P(\chi^2 > X^2) = 1 - p\text{chisq}(X^2, n-1)
\]

\[
\text{one-sided-less: } p = P(\chi^2 < X^2) = p\text{chisq}(X^2, n-1)
\]

• Critical values: \( x_\alpha \) has right-tail probability \( \alpha \)

\[
P(\chi^2 > x_\alpha | H_0) = \alpha \iff x_\alpha = q\text{chisq}(1 - \alpha, n - 1).
\]

• Rejection regions: let \( \alpha \) be the significance.
  \( x_\alpha \) has right-tail probability \( \alpha \)

  Right-sided rejection region: \([x_\alpha, \infty)\)
  Left-sided rejection region: \((-\infty, x_{1-\alpha}]\)
  Two-sided rejection region: \((-\infty, x_{1-\alpha/2}] \cup [x_{\alpha/2}, \infty)\)

\(^2\) test for goodness of fit for categorical data

• Use: Test whether discrete data fits a specific finite probability mass function.

• Data: An observed count \( O_i \) in cell \( i \) of a table.

• Assumptions: None

• \( H_0 \): The data was drawn from a specific discrete distribution.

• \( H_A \): The data was drawn from a different distribution

• Test statistic: The data consists of observed counts \( O_i \) for each cell. From the null hypothesis probability table we get a set of expected counts \( E_i \). There are two statistics that we can use:

  \[
  \text{Likelihood ratio statistic } G = 2 \sum O_i \ln \left( \frac{O_i}{E_i} \right)
  \]

  \[
  \text{Pearson’s chi-square statistic } X^2 = \sum \frac{(O_i - E_i)^2}{E_i}.
  \]

It is a theorem that under the null hypothesis \( X^2 \approx G \) and both are approximately chi-square. Before computers, \( X^2 \) was used because it was easier to compute. Now, it is better to use \( G \) although you will still see \( X^2 \) used quite often.
• Degrees of freedom $df$: The number of cell counts that can be freely specified. In the case above, of the $n$ cells $n - 1$ can be freely specified and the last must be set to make the correct total. So we have $df = n - 1$ degrees of freedom.

In other chi-square tests there can be more relations between the cell counts of $df$ might be different from $n - 1$.

• Rule of thumb: Combine cells until the expected count in each cell is at least 5.

• Null distribution: Assuming $H_0$, both statistics (approximately) follow a chi-square distribution with $df$ degrees of freedom. That is both $\phi(G \mid H_0)$ and $\phi(X^2 \mid H_0)$ have the approximately same pdf as $Y \sim \chi^2(df)$.

• $p$-value:
  \[ p = P(Y > G) = 1 - \text{pchisq}(G, df) \]
  \[ p = P(Y > X^2) = 1 - \text{pchisq}(X^2, df) \]

• Critical values: $c_\alpha$ has right-tail probability $\alpha$
  \[ P(Y > c_\alpha \mid H_0) = \alpha \iff c_\alpha = q\text{chisq}(1 - \alpha, df). \]

• Rejection regions: let $\alpha$ be the significance.
  We expect $X^2$ to be small if the fit of the data to the hypothesized distribution is good. So we only use a right-sided rejection region: $[c_\alpha, \infty)$.

One-way ANOVA ($F$-test for equal means)

• Use: Compare the data means from $n$ groups with $m$ data points in each group.

• Data:
  \[
x_{1,1}, x_{1,2}, \ldots, x_{1,m} \\
x_{2,1}, x_{2,2}, \ldots, x_{2,m} \\
\vdots \\
x_{n,1}, x_{n,2}, \ldots, x_{n,m}
\]

• Assumptions: Data for each group is an independent normal sample drawn from distributions with (possibly) different means but the same variance:
  \[
x_{1,j} \sim N(\mu_1, \sigma^2) \\
x_{2,j} \sim N(\mu_2, \sigma^2) \\
\vdots \\
x_{n,j} \sim N(\mu_n, \sigma^2)
\]

The group means $\mu_i$ are unknown and possibly different. The variance $\sigma$ is unknown, but the same for all groups.

• $H_0$: All the means are identical $\mu_1 = \mu_2 = \ldots = \mu_n$.

• $H_A$: Not all the means are the same.

• Test statistic: $f = \frac{MS_B}{MS_W}$, where
\[ \bar{x}_i = \text{mean of group } i = \frac{x_{i,1} + x_{i,2} + \ldots + x_{i,m}}{m}. \]
\[ \overline{x} = \text{grand mean of all the data}. \]
\[ s_i^2 = \text{sample variance of group } i = \frac{1}{m-1} \sum_{j=1}^{m} (x_{i,j} - \bar{x}_i)^2. \]
\[ MS_B = \text{between group variance} = m \times \text{sample variance of group means} = \frac{m}{n-1} \sum_{i=1}^{n} (\bar{x}_i - \overline{x})^2. \]
\[ MS_W = \text{average within group variance} = \frac{s_1^2 + s_2^2 + \ldots + s_n^2}{n}. \]

- Idea: If the \( \mu_i \) are all equal, this ratio should be near 1. If they are not equal then \( MS_B \) should be larger while \( MS_W \) should remain about the same, so \( f \) should be larger. We won’t give a proof of this.
- Null distribution: \( \phi(f \mid H_0) \) is the pdf of \( F \sim F(n-1, n(m-1)). \) This is the \( F \)-distribution with \( n-1 \) and \( n(m-1) \) degrees of freedom. Several \( F \)-distributions are plotted below.
- \( p \)-value: \( p = P(F > f) = 1 - \text{pf}(f, n-1, n*(m-1)) \)

**Notes:**
1. ANOVA tests whether all the means are the same. It does not test whether some subset of the means are the same.
2. There is a test where the variances are not assumed equal.
3. There is a test where the groups don’t all have the same number of samples.

**\( F \)-test for equal variances**

- Use: Compare the variances from two groups.
- Data: \( x_1, x_2, \ldots, x_n \) and \( y_1, y_2, \ldots, y_m. \)
- Assumptions: Both groups of data are independent normal samples:
  \[ x_i \sim N(\mu_x, \sigma_x^2) \]
  \[ y_j \sim N(\mu_y, \sigma_y^2) \]
where $\mu_x$, $\mu_y$, $\sigma_x$, and $\sigma_y$ are all unknown.

- $H_0$: $\sigma_x = \sigma_y$
- $H_A$:
  - Two-sided: $\sigma_x \neq \sigma_y$
  - One-sided-greater: $\sigma_x > \sigma_y$
  - One-sided-less: $\sigma_x < \sigma_y$

- Test statistic: $f = \frac{s^2_x}{s^2_y}$
  where $s^2_x$ and $s^2_y$ are the sample variances of the data.
- Null distribution: $\phi(f \mid H_0)$ is the pdf of $F \sim F(n-1, m-1)$.
  ($F$-distribution with $n - 1$ and $m - 1$ degrees of freedom.)
- $p$-value:
  - Two-sided: $p = 2 \times \min(pf(f, n-1, m-1), 1 - pf(f, n-1, m-1))$
  - One-sided-greater: $p = P(F > f) = 1 - pf(f, n-1, m-1)$
  - One-sided-less: $p = P(F < f) = pf(f, n-1, m-1)$
- Critical values: $f_\alpha$ has right-tail probability $\alpha$
  $P(F > f_\alpha \mid H_0) = \alpha \iff f_\alpha = qf(1 - \alpha, n - 1, m - 1)$. 