Inferences about central values ($\boldsymbol{\mu}$)

$Y \sim$ normal(μ, σ^2)

Inferences about μ using data: $y_1, y_2, ..., y_n$ (collected as a random sample)

Point estimate

$$\widehat{\mu} = \overline{y} = \frac{1}{n}(y_1 + y_2 + \cdots + y_n)$$

How good is the estimate?

Confidence interval (σ^2 known)

A $100(1-\alpha)\%$ confidence interval (CI) for μ is

$$\left(\overline{y} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \quad \overline{y} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = (l, u)$$

A CI is constructed so that under repeated sampling, the long-run proportion of such intervals that *contain* μ will be $(1-\alpha)$:

$$\mathrm{P}\left(\overline{Y} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{Y} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$\left(\text{What is } \mathrm{P}\left(\overline{y} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{y} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)?\right)$

What does the estimate say about contending models?

Statistical hypothesis tests (σ^2 known)

First model: $\mu = \mu_0$ (known constant)

"Null hypothesis" H_0 (often the skeptic's hypothesis)

Second model: $\mu > \mu_0$

"Alternative hypothesis" H_a (often the research hypothesis)

Distribution of $\overline{Y}$ under each model:

Decision to abandon first model (null) for the second model (alternative) based on a **rejection region**. Reject H_0 in favor of H_a if $\overline{y}$ exceeds a **critical value**:

$$\text{reject } H_0 \text{ if } \overline{y} \geq y_a$$

$$\text{do not reject } H_0 \text{ if } \overline{y} < y_a$$

α: probability of rejecting the null model, given the null model is the true model (set small enough to convince skeptics)

$\alpha = 0.05$: everyday science
$\alpha = 0.01$: important stuff (medical, etc.)

Hypotheses: H_0: $\mu \leq \mu_0$ (= , but get all $<$ values free!)

H_a: $\mu > \mu_0$

Test statistic: rephrase in terms of z (instead of $\overline{y}$)

$$z = \frac{\overline{y} - \mu_0}{(\sigma/\sqrt{n})}$$

Rejection region (decision rule)

reject H_0 if $z \geq z_\alpha$

do not reject H_0 if $z < z_\alpha$

$$\left(y_\alpha = \frac{\sigma}{\sqrt{n}} z_\alpha + \mu_0 \right)$$

One-sided in the other direction (left-tailed test)

Hypotheses: H_0: $\mu \geq \mu_0$

H_a: $\mu < \mu_0$

Rejection region (decision rule)

reject H_0 if $z \leq -z_\alpha$

do not reject H_0 if $z > -z_\alpha$

Two-sided hypothesis test

H_0: $\mu = \mu_0$

H_a: $\mu \neq \mu_0$

Rejection region: reject H_0 if $\overline{y}$ is outside of $\left(y_{1-\frac{\alpha}{2}}, y_{\frac{\alpha}{2}}\right)$

Test statistic: $z = \frac{\overline{y} - \mu_0}{(\sigma/\sqrt{n})}$

Rejection region:

reject H_0 if $|z| \geq z_{\alpha/2}$

do not reject H_0 if $|z| < z_{\alpha/2}$

P-values for hypothesis tests

The ***p*-value** for a hypothesis test is the probability that, under repeated sampling, the test statistic would be *as extreme* as the observed value of the test statistic, given that the null hypothesis is true.

Right-sided: H_0: $\mu \leq \mu_0$

H_a: $\mu > \mu_0$

$p = P(Z \geq z)$

reject H_0 if $z \geq z_\alpha \Leftrightarrow$ reject H_0 if $p \leq \alpha$

Left-sided: H_0: $\mu \geq \mu_0$

H_a: $\mu < \mu_0$

$p = P(Z \leq z)$

reject H_0 if $z \leq -z_\alpha \Leftrightarrow$ reject H_0 if $p \leq \alpha$

Two-sided: H_0: $\mu = \mu_0$

H_a: $\mu \neq \mu_0$

$p = 2P(Z \geq |z|)$

reject H_0 if $|z| \geq z_{\alpha/2} \Leftrightarrow$ reject H_0 if $p \leq \alpha$

Note: the p-value is *not* the probability of the null hypothesis!! ("probability of a hypothesis" is meaningless in "frequentist" statistics)

Relationship between CIs and two-sided hypothesis tests

A $100(1-\alpha)\%$ CI for μ is the set of all values of μ_0 for which the null hypothesis H_0: $\mu = \mu_0$ would not be rejected in a test against the alternative hypothesis H_a: $\mu \neq \mu_0$

Inferences for μ with σ^2 unknown

$Y \sim$ normal(μ, σ^2); $Y_1, Y_2, ..., Y_n$ a random sample; then $T = \frac{\overline{Y}-\mu}{(S/\sqrt{n})}$ has a **Student's t distribution** with $n - 1$ **degrees of freedom** (df).

Student's t distribution approaches a standard normal distribution as $n \to \infty$. Table 2 (Appendix, p. 1093) lists selected percentiles of the Student's t distribution for various df values. The $100(1-\alpha)$th percentile of the Student's t distribution is denoted as t_a (with some particular df specified)

Confidence interval (σ^2 unknown)

A $100(1-\alpha)\%$ CI for μ is

$$\left(\overline{y} - t_{\alpha/2}\frac{s}{\sqrt{n}}, \quad \overline{y} + t_{\alpha/2}\frac{s}{\sqrt{n}}\right) = (l, u)$$

(df $= n - 1$)

Hypothesis tests (σ^2 unknown)

Case 1. H_0: $\mu \leq \mu_0$ vs. H_a: $\mu \geq \mu_0$ (right-tailed test)

Case 2. H_0: $\mu \geq \mu_0$ vs. H_a: $\mu \leq \mu_0$ (left-tailed test)

Case 3. H_0: $\mu = \mu_0$ vs. H_a: $\mu \neq \mu_0$ (two-tailed test)

α specified; $\mathrm{df} = n - 1$

Test statistic:

$$t = \frac{\overline{y} - \mu_0}{\left(s/\sqrt{n}\right)}$$

p-value:

Case 1. $p = \mathrm{P}(T \geq t)$

Case 1. $p = \mathrm{P}(T \leq t)$

Case 1. $p = 2 \cdot \mathrm{P}(T \geq |t|)$

Rejection region:

Case 1. Reject H_0 if $t \geq t_\alpha$ ($p \leq \alpha$)

Case 2. Reject H_0 if $t \leq -t_\alpha$ ($p \leq \alpha$)

Case 3. Reject H_0 if $|t| \geq t_{\alpha/2}$ ($p \leq \alpha$)