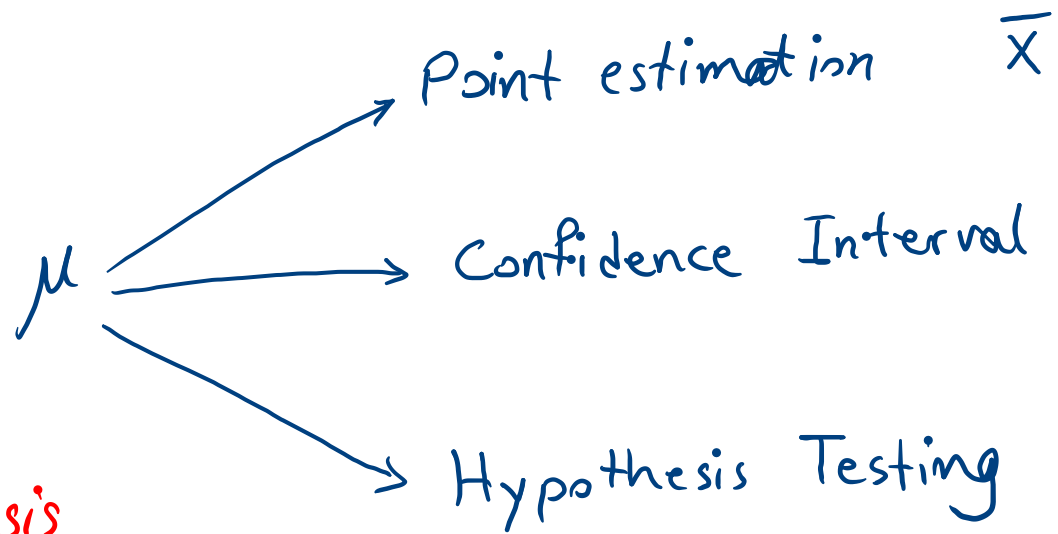


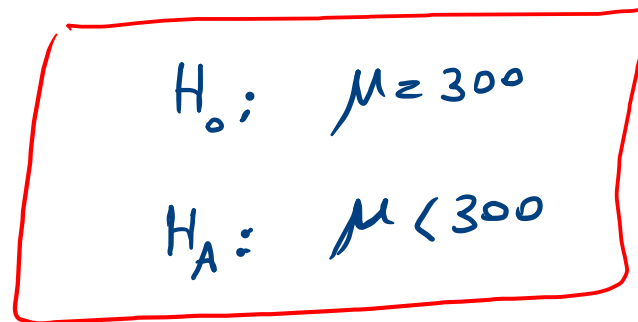
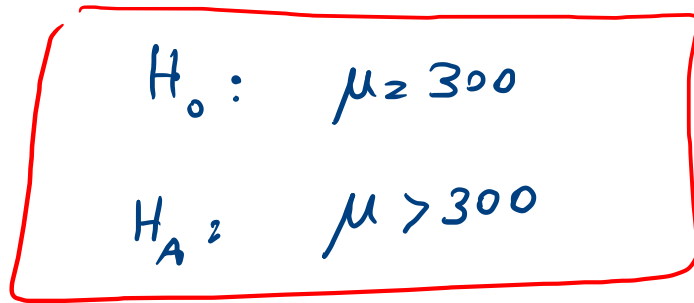
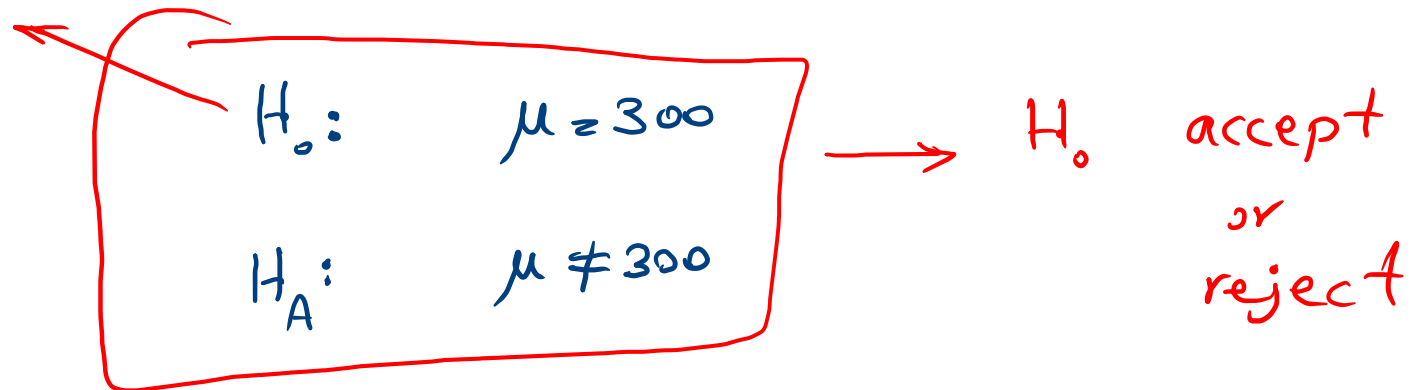
Hypothesis Testing



$$\bar{X} - a \leq \mu \leq \bar{X} + a$$

$1 - \alpha$

Null Hypothesis



آزمون فرض (Hypothesis Testing)

○ آزمون فرض روشی برای بررسی ادعاها و یا فرضیات درباره پارامترهای توزیع در جوامع آماری است.

○ فرض کنید یکی از اساتید دانشگاه تهران در مصاحبه با مطبوعات ادعا می کند که دانشجویان این دانشگاه به طور متوسط در سال ۳۰۰ هزار تومان کتاب می خرند. می خواهیم با استفاده از یک نمونه جمع آوری شده، صحت ادعای ایشان را بررسی کنیم.

○ فرض کنید: $\bar{X} = 320$

فرض مقابل
ادعای استاد صحیح نیست
ادعای استاد درباره میانگین صحیح نیست و نمونه مشاهده شده نمی تواند تصادفی باشد.

$$H_A: \mu \neq 300$$

فرض صفر
ادعای استاد صحیح است
ادعای استاد درباره میانگین صحیح است و نمونه جمع آوری شده به طور تصادفی دارای میانگین ۳۲۰ شده است.

$$H_0: \mu = 300$$

آزمون فرض



○ آزمون فرض شباهت زیادی به یک دادگاه دارد:

○ فرض صفر (H_0): متهم بیگناه است

○ فرض مقابل (H_A): متهم گناهکار است

○ شواهدی ارائه می شود:

○ جمع آوری داده

○ آیا در صورت صحیح بودن فرض صفر، امکان داشت که داده مشاهده شده به طور تصادفی اتفاق افتاده باشد؟

○ **بله**: نمی توانیم فرض H_0 را رد کنیم.

○ **خیر**: فرض H_0 را رد می کنیم.

چارچوب آزمون فرض

○ با یک فرض صفر (null hypothesis) که ادعای مورد بحث را نمایش می‌دهد شروع می‌کنیم. برای فرض H_0 همیشه از نماد $=$ استفاده می‌کنیم.

○ سپس فرض مقابل (alternative hypothesis) را مطرح می‌کنیم که سوال تحقیق را بیان می‌کند، به عبارت دیگر فرضی که به دنبال آزمایش آن هستیم. برای فرض H_A از یکی از نمادهای $<, >, \neq$ استفاده می‌کنیم.

○ آزمون فرض را به کمک قضیه حد مرکزی و با فرض درست بودن فرض صفر اجرا می‌کنیم.

مثال

○ یک نمونه ۵۰ تایی از دانشجویان دانشگاه تهران انتخاب و از آنها سوال شده است که در سال چه مبلغی صرف خرید کتاب می کنند. پاسخ این دانشجویان دارای متوسط ۳۲۰ هزار تومان با انحراف معیار ۱۷۴ هزار تومان است.

○ یکی از اساتید دانشگاه تهران در مصاحبه با مطبوعات ادعا می کند که دانشجویان این دانشگاه به طور متوسط در سال ۳۰۰ هزار تومان کتاب میخرند.

• **فرض صفر:** دانشجویان به طور متوسط ۳۰۰ هزار تومان کتاب میخرند. $H_0: \mu = 300$

• **فرض مقابل:** دانشجویان به طور متوسط بیش از ۳۰۰ هزار تومان کتاب میخرند. $H_A: \mu > 300$

فرضیات همیشه درباره
پارامتر جامعه هستند.

اجرای آزمون فرض به کمک بازه اطمینان

○ برای اجرای آزمون فرض دو راه کلی وجود دارد:

← ○ آزمون فرض با استفاده از بازه اطمینان

← ○ آزمون فرض با استفاده از p-value

○ آزمون فرض به کمک بازه اطمینان:

○ اگر فرض صفر داخل بازه اطمینان ۹۵٪ قرار بگیرد، نمی‌توانیم آن را رد کنیم و در غیر

این صورت آن را رد می‌کنیم.

○ این روش سریع است ولی میزان خطای آزمون را مشخص نمی‌کند.

مثال

$$1 - \alpha \approx 1 \Rightarrow \alpha \approx 0 \rightarrow Z_{1 - \frac{\alpha}{2}}$$

فرض صفر: دانشجویان به طور متوسط ۳۰۰ هزار تومان کتاب میخرند.
 $H_0: \mu = 300$

فرض مقابل: دانشجویان به طور متوسط بیش از ۳۰۰ هزار تومان کتاب میخرند.
 $H_A: \mu > 300$

○ یک نمونه ۵۰ تایی از دانشجویان دانشگاه تهران انتخاب و از آنها سوال شده است که در سال چه مبلغی صرف خرید کتاب می کنند. پاسخ این دانشجویان دارای متوسط ۳۲۰ هزار تومان با انحراف معیار ۱۷۴ هزار تومان است. دیدیم که بازه اطمینان ۹۵٪ به کمک این نمونه جمع آوری شده برابر است با: (272, 368)



○ از آنجا که فرض صفر در بازه اطمینان قرار دارد، نمی توانیم آن را رد کنیم.

آزمون فرض به روش p-value

- احتمال مشاهده خروجی نمونه یا مفرطتر به شرط درست بودن فرض صفر را p-value می‌نامیم:

$$\text{p-value} = P(\underbrace{\bar{X} = 320}_{\text{observed}} \text{ or more extreme outcome} \mid \underbrace{\bar{X} > 320}_{\text{H}_0 \text{ true}})$$

- برای مثال:

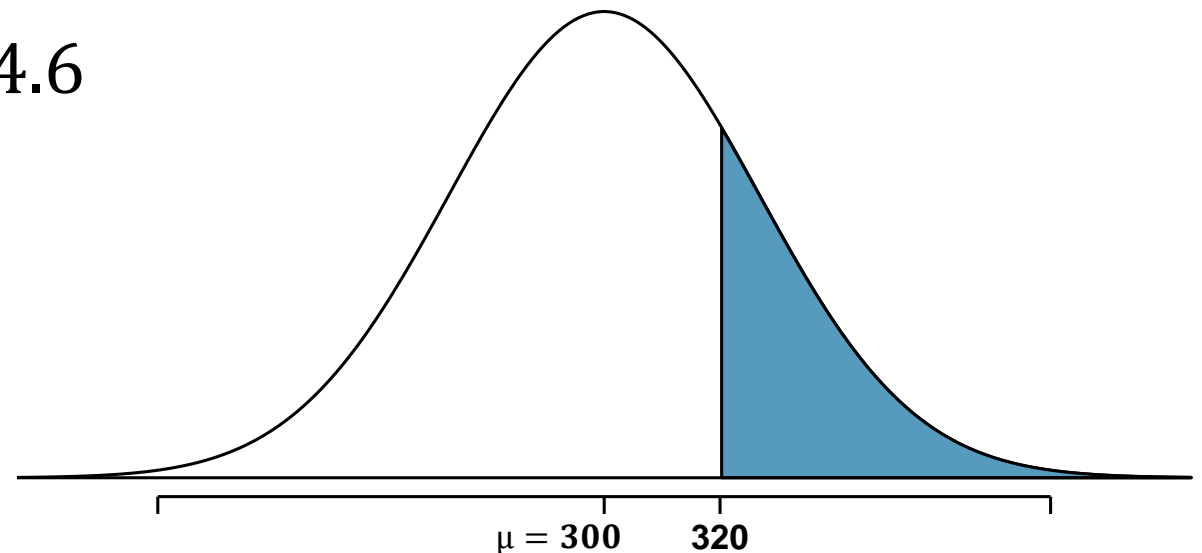
$$\text{p-value} = P(\bar{X} > 320 \mid H_0: \mu = 300)$$

$$S = 174, n = 50 \Rightarrow S/\sqrt{n} = 24.6$$

$$\bar{X} \sim N(\mu = 300, S/\sqrt{n} = 24.6)$$

$$\text{آماره آزمون: } Z = \frac{320 - 300}{24.6} = 0.81$$

$$\text{p-value} = P(Z > 0.81) = 0.209$$

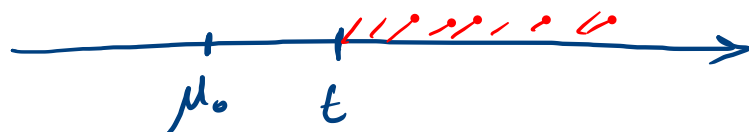


$$H_0: \mu = \mu_0$$

$$H_A: \mu > \mu_0$$

$$p\text{-value} = P(\bar{x} > t \mid \mu = \mu_0)$$

$$\bar{x} = t$$



$$H_0: \mu = \mu_0$$

$$H_A: \mu < \mu_0$$

$$p\text{-value} = P(\underbrace{\bar{x} < t}_{\bar{x} - t < 0} \mid \mu = \mu_0)$$

$$\bar{x} - t < 0$$

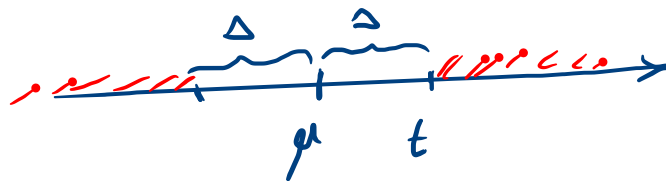


$$H_0: \mu = \mu_0$$

$$|t - \mu_0| = \Delta$$

$$H_A: \mu \neq \mu_0$$

$$p\text{-value} = P(\bar{x} < \mu - \Delta \text{ or } \bar{x} > \mu + \Delta \mid \mu = \mu_0)$$



Hypothesis testing

$H_0: \mu = \mu_0 \rightarrow$ null hypothesis

$H_A: \mu \neq \mu_0$
 $>$
 $<$

\bar{x}

① Confidence interval

② p-value

p-value = $P(\text{observed value or more extreme value} \mid H_0 \text{ is true})$

$$\bar{X} = 320$$

$$H_0: \mu = 300$$

$$H_A: \mu > 300$$

$$p\text{-value} = P(\bar{X} \geq 320 \mid \mu = 300)$$

one-sided
hypothesis
testing

two-sided hypothesis testing

$$\bar{X} = 320$$

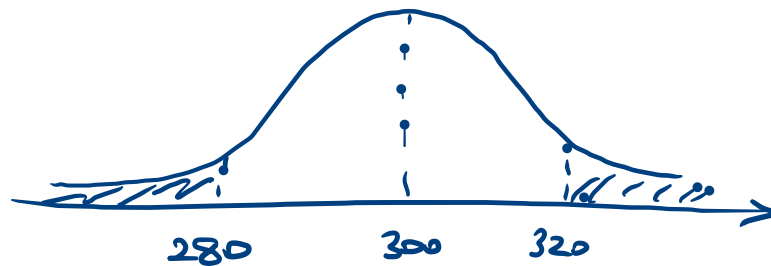
$$H_0: \mu = 300$$

$$H_A: \mu \neq 300$$

$$p\text{-value} = P(\bar{X} \geq 320 \text{ or } \bar{X} \leq 280 \mid \mu = 300)$$

$$= P(\bar{X} \geq 320 \mid \mu = 300) + P(\bar{X} \leq 280 \mid \mu = 300)$$

$$\bar{X} \sim \mathcal{N}\left(300, \frac{\sigma^2}{n}\right)$$



آزمون فرض به روش p-value

$$H_0: \mu = 300$$

$$H_A: \mu > 300$$

- احتمال مشاهده خروجی نمونه یا مفرطتر به شرط درست بودن فرض صفر را p-value می‌نامیم:

$$\text{p-value} = P(\text{observed or more extreme outcome} \mid H_0 \text{ true})$$

- برای مثال:

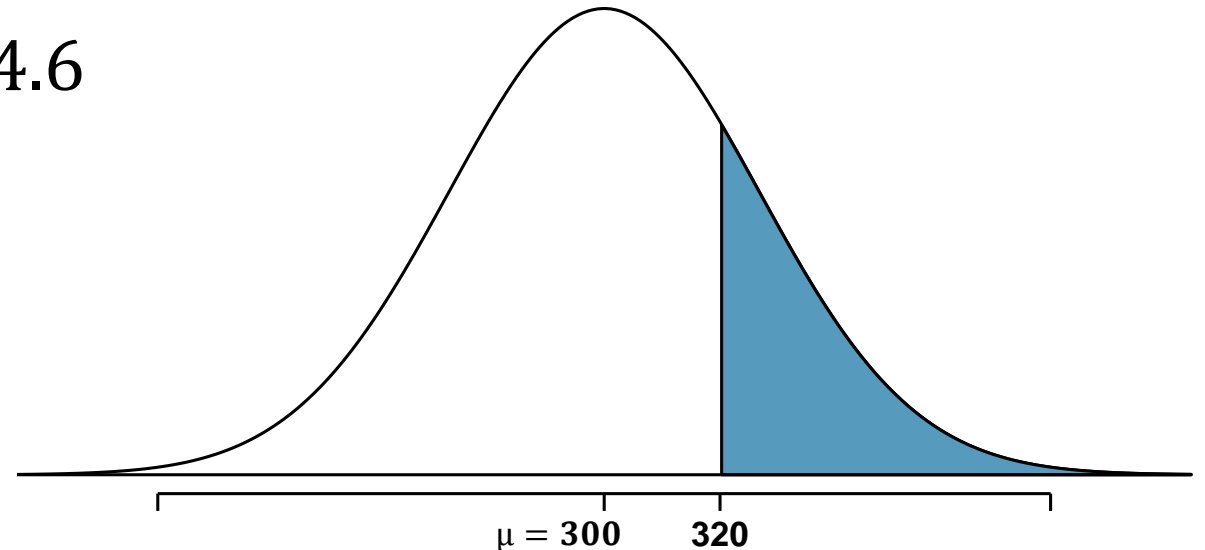
$$\text{p-value} = P(\bar{X} \geq 320 \mid H_0: \mu = 300)$$

$$S = 174, n = 50 \Rightarrow S/\sqrt{n} = 24.6$$

$$\rightarrow \bar{X} \sim N(\mu = 300, S/\sqrt{n} = 24.6)$$

$$\text{آماره آزمون: } Z = \frac{320 - 300}{24.6} = 0.81$$

$$\text{p-value} = P(Z > 0.81) = 0.209$$

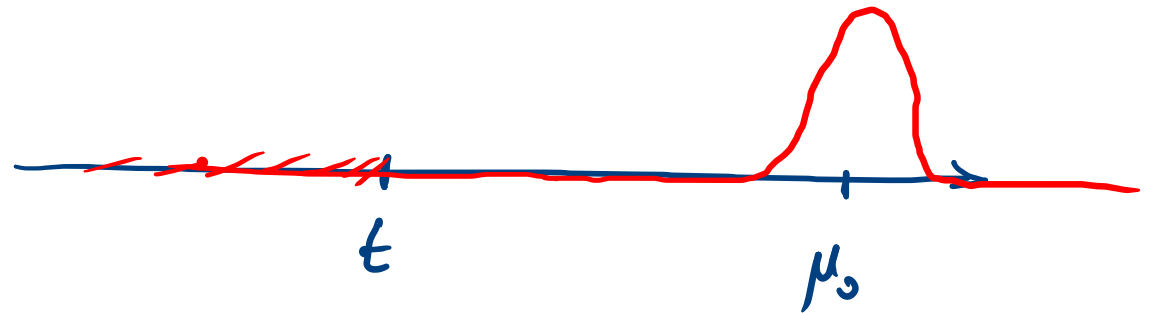


$$p\text{-value} = P(\bar{X} \geq 320 \mid \mu = 300) = P\left(\frac{\bar{X} - \mu}{s/\sqrt{n}} \geq \frac{320 - 300}{24.6}\right)$$

Z

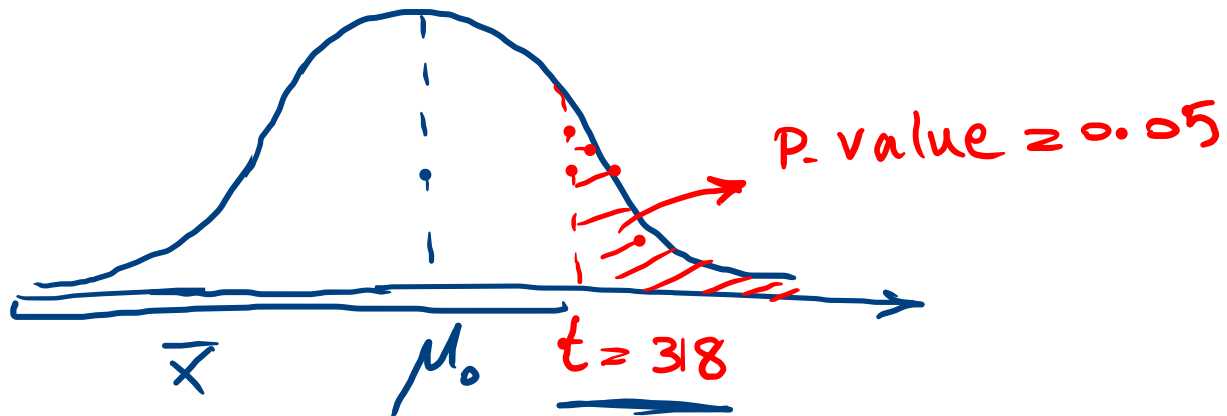
$$= P(Z \geq 0.81) = 1 - \Phi(0.81) = 0.209 > 0.05 \rightarrow H_0 \text{ accept}$$

$$\max_{\mu_0} P(\bar{X} \geq t \mid \mu = \mu_0)$$



$$p\text{-value} = P(\overline{X} \geq t \mid H_0 \text{ is true}) = 1 - \Phi\left(\frac{t - \mu_0}{\frac{\sigma}{\sqrt{n}}}\right) = 0.05$$

318



$$H_A: \mu > \mu_0$$

تصمیم‌گیری بر مبنای p-value

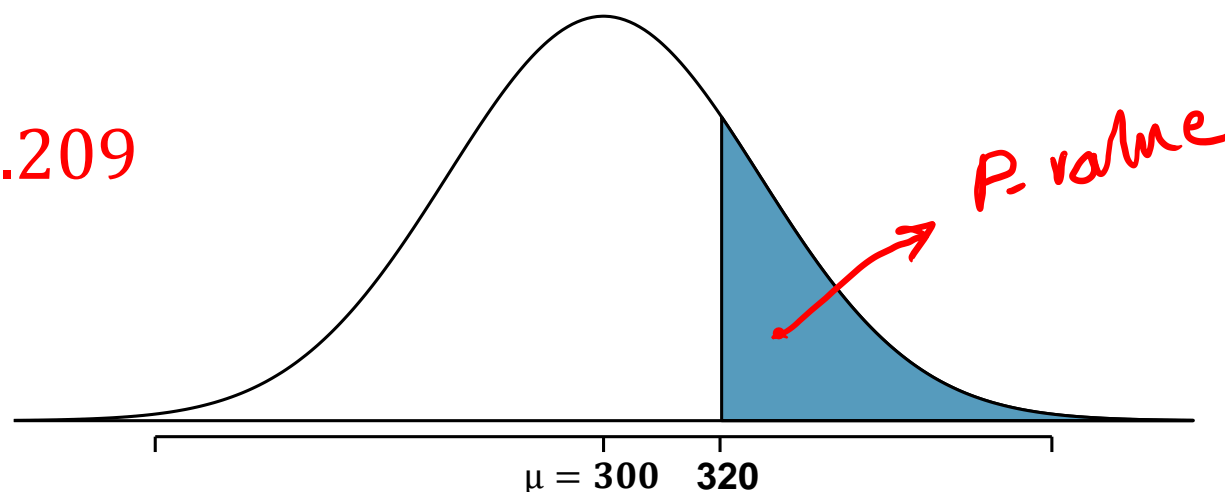
○ اگر p-value کوچک باشد (کمتر از سطح اهمیت (significance level) یا α که معمولاً ۰.۰۵٪ در نظر گرفته می‌شود)، فرض H_0 رد می‌شود.

○ اما اگر p-value بزرگتر از ۰.۰۵٪ باشد فرض H_0 رد نمی‌شود.

$$\text{p-value} = P(\bar{X} > 320 | \mu = 300) = 0.209$$

$$\text{significance level: } \alpha = 0.05$$

$$\text{p-value} = 0.209 > \underline{\underline{0.05}}$$



از آنجا که p-value بزرگ است، ما نمی‌توانیم H_0 را رد کنیم.

آزمون فرض دوطرفه

○ در بسیاری از مواقع سوال تحقیق ما تنها به انحراف از فرض صفر در یک جهت مربوط نمی‌شود.

○ به عبارت دیگر فرض مقابل به صورت $(\mu <)$ و یا $(\mu >)$ مطرح نمی‌شود، بلکه به صورت $(\mu \neq)$ مطرح می‌شود.

○ چنین آزمونی را آزمون فرض دوطرفه (two-sided) می‌نامیم.

○ تعریف p-value در این حالت مشابه قبل است، اما نحوه محاسبه آن به دلیل این که احتمال مفرط‌تر بودن از هر دو طرف باید در نظر گرفته شود متفاوت است.

مثال

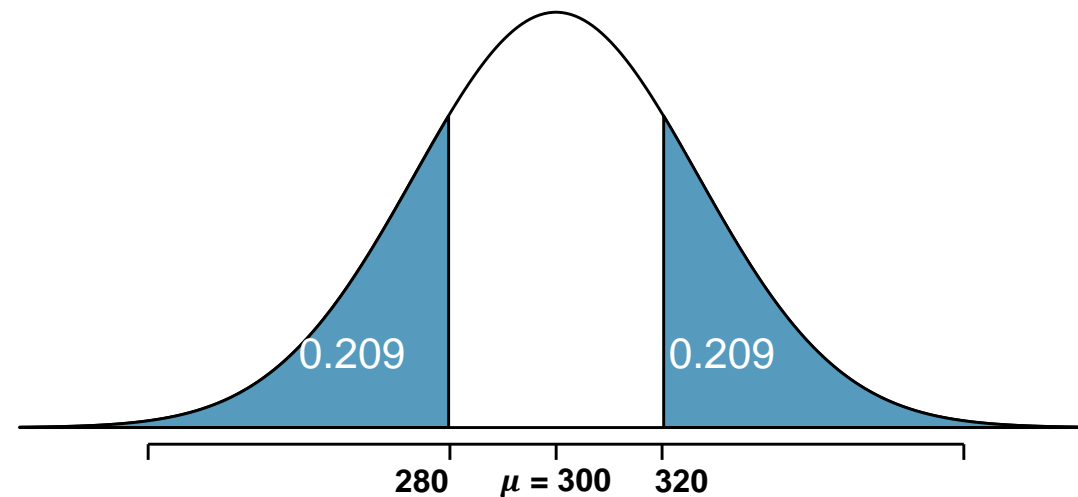
$$H_0: \mu = 300$$

فرض صفر: دانشجویان به طور متوسط ۳۰۰ هزار تومان کتاب میخرند.

$$\rightarrow H_A: \mu \neq 300$$

فرض مقابل: دانشجویان به طور متوسط بیشتر یا کمتر از ۳۰۰ هزار تومان کتاب میخرند.

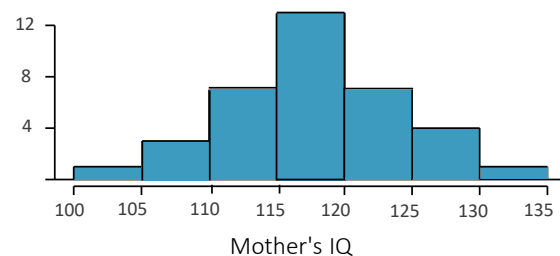
$$\text{p-value} = P(\bar{X} > 320 \text{ or } \bar{X} < 280 | H_0: \mu = 300)$$



$$\text{p-value} = P(Z > 0.81) + P(Z < -0.81) = \underbrace{0.209} + \underbrace{0.209} = 0.418$$

مثال ۲

- گروهی از محققین که بر روی ویژگی‌های کودکان با استعداد مطالعه می‌کنند، نمونه‌ای شامل ۳۶ کودک با استعداد ۴ ساله را از یک شهر بزرگ جمع‌آوری کرده‌اند. در این مطالعه ضریب هوشی مادران این کودکان اندازه‌گیری شده است که در نمودار زیر نتایج آن را مشاهده میکنید. می‌دانیم متوسط ضریب هوشی متوسط مردم این شهر برابر با ۱۰۰ است. به کمک آزمون فرض بررسی کنید که آیا ضریب هوشی مادران کودکان با استعداد با میانگین جامعه یکسان است یا خیر؟



n	36
min	101
mean	118.2
sd	6.5
max	131

$$H_0: \mu = 100$$

$$H_A: \mu \neq 100$$

$$P\text{-value} = P(\bar{X} \geq 118.2 \text{ or } \bar{X} \leq 81.8 \mid \mu = 100)$$

$$= 2 P(\bar{X} \geq 118.2 \mid \mu = 100)$$

$$= 2 P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \geq \frac{118.2 - 100}{\frac{6.5}{\sqrt{36}}}\right) = 2 P(Z \geq 18) \approx 0 < 0.05$$

h_0 is rejected

test statistic

z-score

آزمون فرض برای نسبت

$$H_0 : p = \text{null value}$$

(۱) برپایی آزمون فرض با توجه به مساله:

$$H_A : p < \text{or } > \text{or } \neq \text{null value}$$

(۲) محاسبه تخمین نقطه‌ای: \hat{p}

آماره آزمون

(۳) ✓ بررسی شرایط CLT برای نسبت p

(۴) محاسبه آماره آزمون و p-value:

$$\textcircled{Z} = \frac{\textcircled{\hat{p}} - \textcircled{p}}{\sqrt{\textcircled{p(1-p)}}/n}$$

(۵) مقایسه p-value با α و تصمیم‌گیری

مثال

○ در یک نظرسنجی که از ۱۹۸۳ نفر در امریکا انجام شد، باور افراد نسبت به نظریه تکامل مورد سوال قرار گرفت. ۶۰٪ افراد شرکت کننده در این نظرسنجی گفته اند که به این نظریه باور دارند. آیا می توان ادعا کرد که اکثریت مردم امریکا نظریه تکامل را قبول دارند؟

$$H_0: p = 0.5 \quad \hat{p} = 0.6 \quad n = 1983$$

$$H_A: p > 0.5$$

$$np > 10$$

$$1983 \times 0.5 > 10$$



$$n(1-p) > 10$$

$$1983 \times (1-0.5) > 10$$

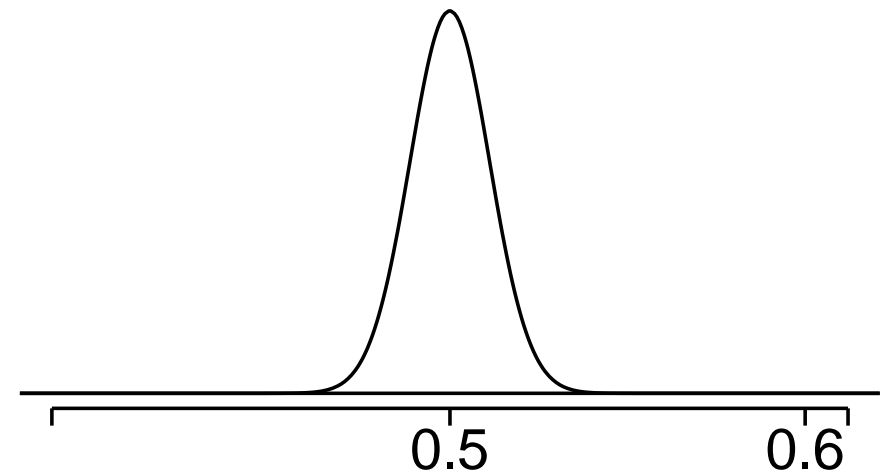
$$Z = \frac{0.6 - 0.5}{\sqrt{0.5 \times (1-0.5) / 1983}}$$

ادامہ مثال

$$\hat{p} \sim N(p = 0.5, \sqrt{\frac{0.5 \times 0.5}{1983}} \approx 0.0112)$$

$$Z = \frac{0.6 - 0.5}{0.0112} \approx 8.92$$

$$\text{p-value} = P(Z > 8.92) \approx 0 \ll 0.05 \rightarrow \text{Reject } H_0$$



مثال ۲

- مدیر کل دانشجویی وزارت بهداشت آمار اعتیاد دانشجویان به مواد مخدر را ۸/۲ درصد اعلام میکند. دانشجویی برای بررسی صحت این ادعا با انتخاب یک نمونه تصادفی ۵۰۰ نفره و انجام تست اعتیاد به جمع‌آوری داده می‌پردازد. ۳۰ نفر از افراد بررسی شده معتاد تشخیص داده می‌شوند. آیا با توجه به داده جمع‌آوری شده، ادعای مدیر وزارت بهداشت صحیح است؟

$$n = 500 \quad \hat{p} = \frac{30}{500} = 0.06$$

$$H_0: p = 0.082$$

$$500 \times 0.082 > 10$$

$$500 \times (1 - 0.082) > 10$$



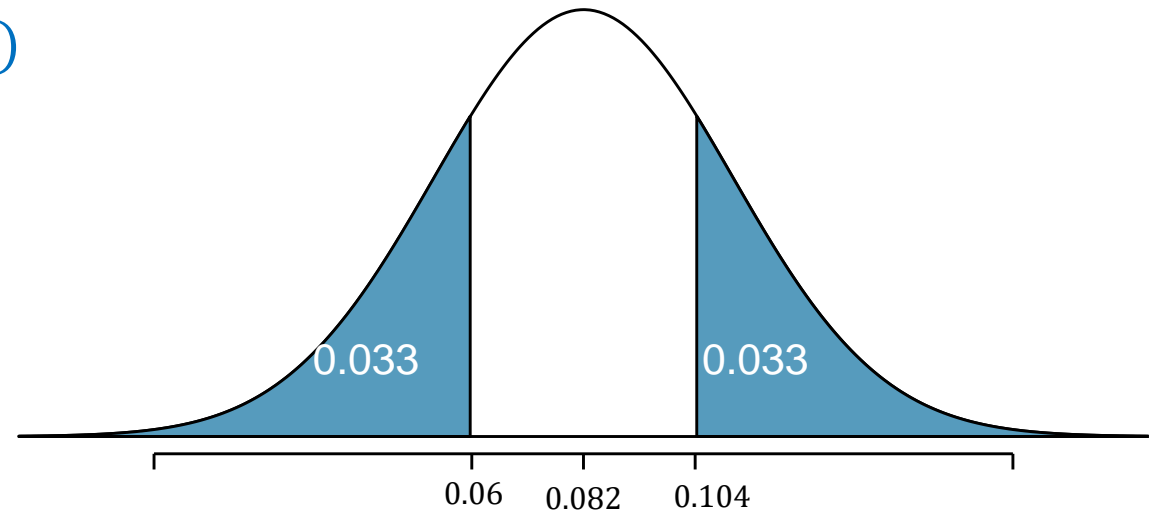
$$H_A: p \neq 0.082$$

$$Z = \frac{0.06 - 0.082}{\sqrt{\frac{0.082(1-0.082)}{500}}}$$

ادامه مثال ۲

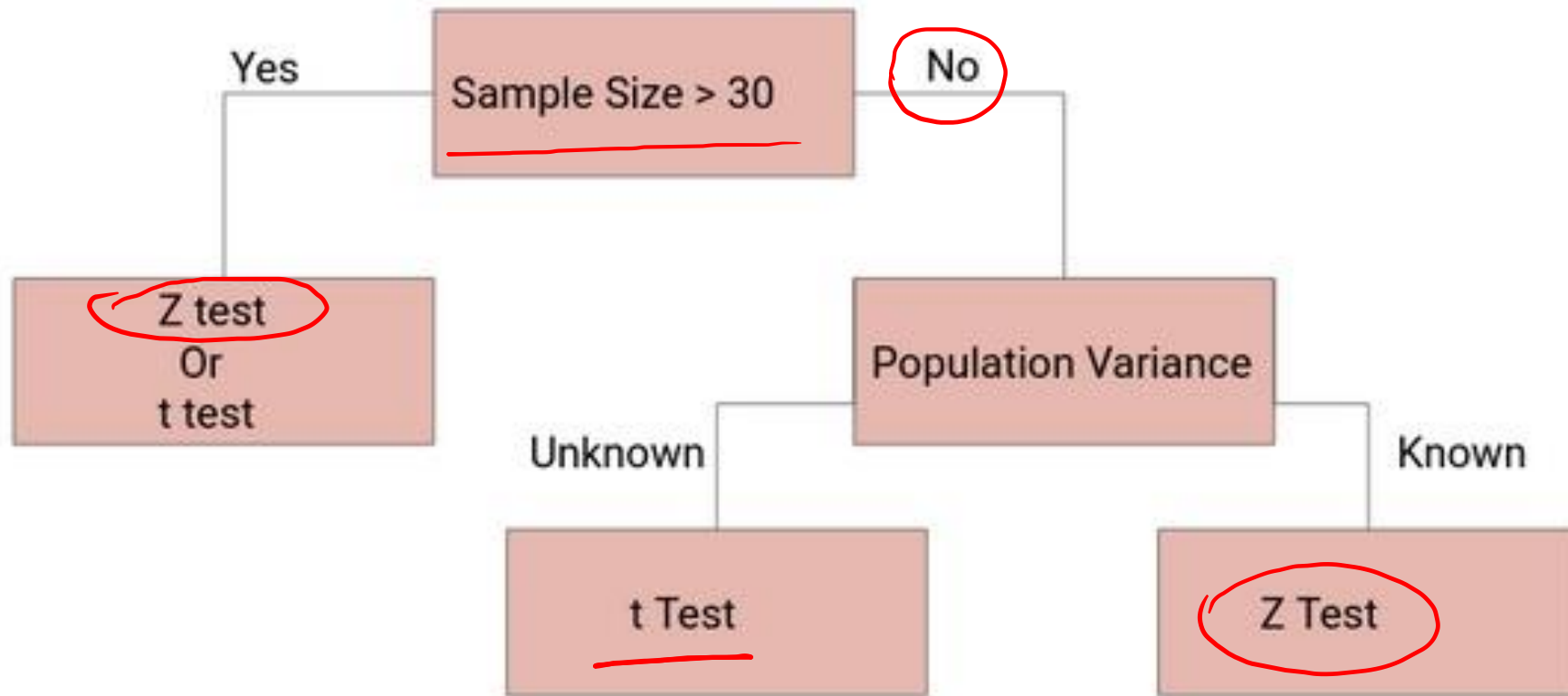
$$\hat{p} \sim N(p = 0.082, \sqrt{\frac{0.082 \times 0.918}{500}} \approx 0.012)$$

$$Z = \frac{0.06 - 0.082}{0.012} \approx \underline{-1.83}$$



$$\text{p-value} = \underbrace{P(Z > 1.83)} + \underbrace{P(Z < -1.83)} \approx 0.066 > 0.05 \rightarrow \text{Fail to Reject } H_0$$

Z-test vs. T-test



Hypothesis Testing Error Types

		Truth about the population	
		<u>H_0 true</u>	<u>H_a true</u>
Decision based on sample	<u>Reject H_0</u>	Type I error	Correct decision
	<u>Accept H_0</u>	Correct decision	Type II error