

عنوان: آزمون تفاضل میانگین های دو جامعه

مدرس: دکتر سمیه عباسی
هیات علمی دانشگاه آزاد اسلامی اصفهان (خوراسگان)

دانشگاه آزاد اسلامی
واحد اصفهان خوراسگان



دانشگاه آزاد اسلامی
واحد رصفهان خوراکان

آزمون فرضیه تفاضل میانگین‌های دو جامعه

استفاده از این آزمون زمانی است که فرضیه محقق در مورد میانگین دو جامعه باشد. یا به عبارت دیگر بخواهد میانگین شاخص مورد بررسی را در دو سازمان مورد مقایسه قرار دهد.

مثال

- میانگین حقوق کارمندان بخش دولتی بیشتر از میانگین حقوق کارمندان بخش خصوصی است؟

- آیا داروی A باعث کاهش فشارخون می شود؟

- شاغلینی که در کلان شهرها زندگی می کنند میزان استرس شغلی بیشتری نسبت به شاغلین سایر شهرها دارند.

- آیا وجدان کاری بعد از افزایش حقوق بیشتر شده است؟

• مرحله اول) تعریف H_0 و H_1

حالت اول

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$$

حالت دوم

$$\begin{cases} H_0 : \mu_1 \geq \mu_2 \\ H_1 : \mu_1 < \mu_2 \end{cases}$$

حالت سوم

$$\begin{cases} H_0 : \mu_1 \leq \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases}$$

نکته: درجه
آزادی در دو
جامعه

$$df = n_1 + n_2 - 2$$

• مرحله دوم) محاسبه آماره آزمون یا ملاک آزمون

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

(۱) دو جامعه نرمال و
انحراف معیارهای آنها
(جامعه ها)
معلوم

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$df > 30$

(۲) دو جامعه نرمال و
انحراف معیارهای آنها
(جامعه ها)
نا معلوم

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$\sigma_1 = \sigma_2$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$df \leq 30$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

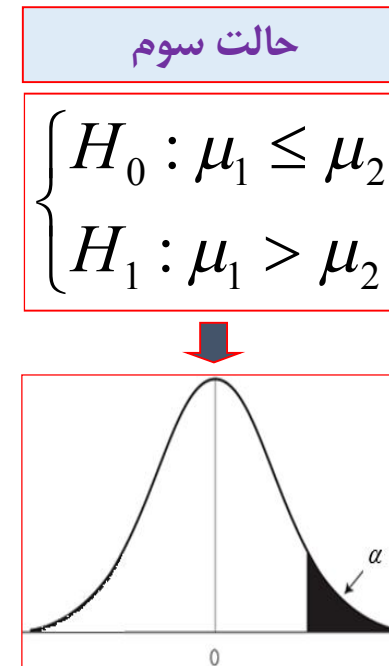
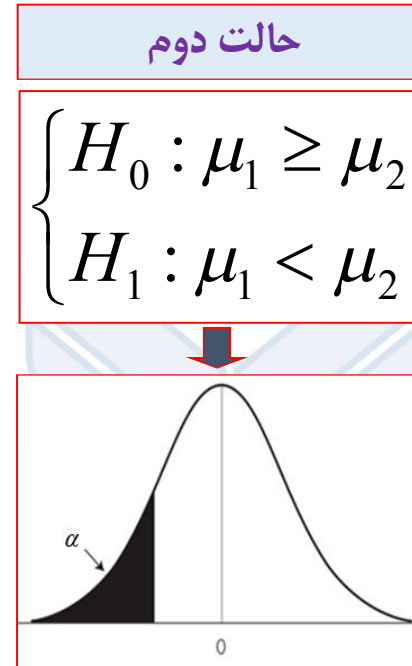
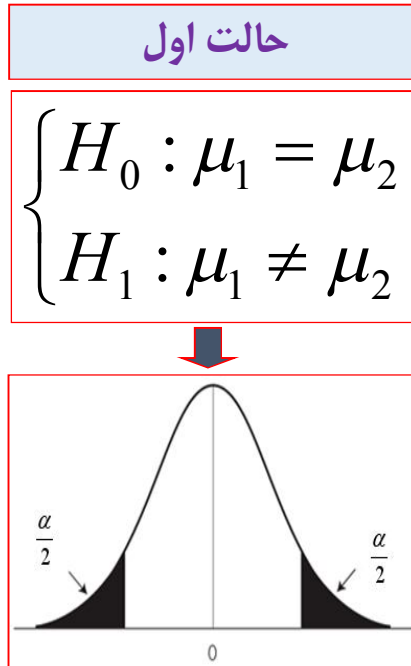
$\sigma_1 \neq \sigma_2$

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

در جوامع غیر نرمال با درجات آزادی بالای ۳۰
آماره Z به کار می رود.

توجه

• مرحله سوم) تعیین ناحیه بحرانی



مرحله چهارم) تصمیم گیری: اگر چنانچه ملاک آزمون در ناحیه بحرانی قرار بگیرد فرضیه H_0 رد می شود.

مثال

فرض کنید در کارخانه‌ای مقایسه عملکرد دو خط تولید مد نظر باشد. برای خط تولید اول (مستهلك) انتظار می‌رود که خروجی آن بر حسب تعداد در روز معادل خط دوم (جدیدتر) باشد. بر حسب تجربیات گذشته واریانس دو خط تولید برابر و مساوی ۴ است. برای این منظور داده‌های تولید ۱۲ روز از خط تولید اول و ۸ روز از خط تولید دوم جمع‌آوری شده است که میانگین در خط تولید اول و دوم $\bar{x}_1 = 11.5$ ، $\bar{x}_2 = 13.5$ ثبت شده است. در سطح خطای ۴ درصد با فرض نرمال بودن فرایند تولید مقوله مورد انتظار را آزمون کنید.

$$n_1 = 12$$

$$n_2 = 8$$

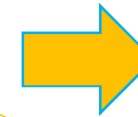
$$\bar{x}_1 = 11.5$$

$$\bar{x}_2 = 13.5$$

$$\sigma_1^2 = \sigma_2^2 = 4$$

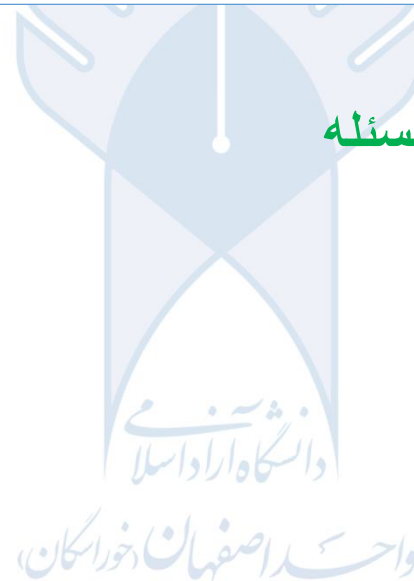
$$\alpha = 0.04$$

ادعای مسئله



$$\mu_1 = \mu_2$$

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$



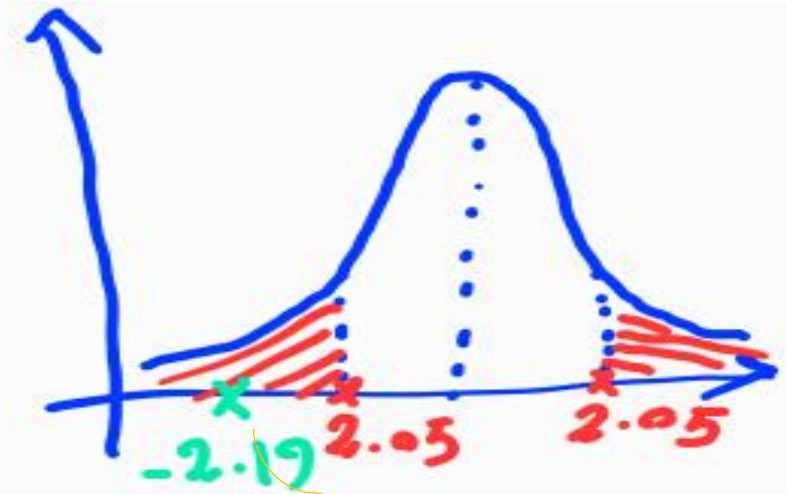
$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\delta_1^2}{n_1} + \frac{\delta_2^2}{n_2}}}$$

$$= \frac{(11.5 - 13.5) - (0)}{\sqrt{\frac{4}{12} + \frac{4}{8}}}$$

$$= -2.19$$

$$\alpha = 0.04 \rightarrow \frac{\alpha}{2} = 0.02$$

$$Z_{\alpha/2} = Z_{0.02} = -2.05$$



مقدار آماره آزمون در ناحیه بحرانی قرار می گیرد. بنابراین، در سطح خطای چهار درصد فرض صفر رد می شود و ادعای مسئله تائید نمی شود و چون آماره آزمون در ناحیه چپ قرار دارد، می توان ادعا کرد که میانگین تولید خط اول که مستهلک تر است کمتر از میانگین تولید خط دوم می باشد.

مثال

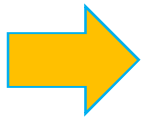
محققى به منظور مقایسه عملکرد مدیران دو سازمان که به روش های مشارکتی و دستوری اداره می شود فرضیه زیر را صورتبندی نموده است:
 “بهره وری سبک مدیریت مشارکتی بهتر از سبک مدیریت دستوری است”
 او بدین منظور از هر سازمان چند نمونه تصادفی انتخاب کرد که اطلاعات آنها بشرح زیر می باشد.
 با فرض برابری واریانسها فرضیه فوق را در سطح ۹۵ درصد آزمون کنید.

	n	\bar{x}	s^2
۱ دستوری	۱۱	۱۲۵	۱۲۱
۲ مشارکتی	۱۰	۱۳۲	۲۲۵

$$\alpha = 0.05$$

$$\sigma_1^2 = \sigma_2^2$$

ادعای مسئله



$$\mu_1 < \mu_2$$

$$\begin{cases} H_0: \mu_1 \geq \mu_2 \\ H_1: \mu_1 < \mu_2 \end{cases}$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$s_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$s_p^2 = \frac{(11-1)121 + (10-1)225}{11+10-2}$$

$$= 170.26$$

$$S_p = \sqrt{S_p^2} = \sqrt{170.26} = 13.04$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

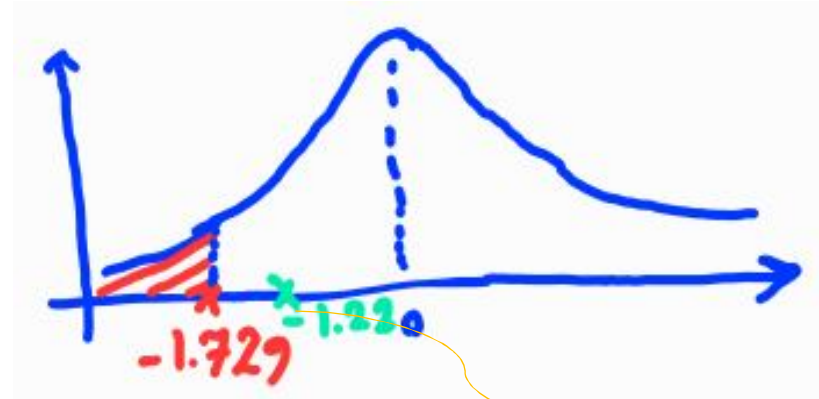
$$t = \frac{(125 - 132) - (0)}{13.04 \sqrt{\frac{1}{11} + \frac{1}{10}}}$$

$$= -1.22$$

$$df = n_1 + n_2 - 2$$

$$= 11 + 10 - 2 = 19$$

$$\alpha = 0.05 \quad t_{0.05, (19)} = 1.729$$



مقدار آماره آزمون در ناحیه پذیرش قرار می گیرد. بنابراین، در سطح خطای پنج درصد فرض صفر پذیرفته و ادعای مسئله در فرض یک تائید نمی شود. لذا می توان گفت بهره وری سبک مدیریت مشارکتی بهتر از مدیریت دستوری نیست.

مثال

فرضیه تحقیقی چنین است: عملکرد مدیران با مهارت ادراکی بالا بیشتر از عملکرد مدیران با مهارت عملی بالاست. میانگین و انحراف معیار عملکرد در ۱۴ مدیر انتخابی با مهارت ادراکی بالا به ترتیب ۴۳ و ۱/۱ و میانگین و انحراف معیار در ۱۵ مدیر با مهارت عملی بالا به ترتیب ۴۰/۷ و ۱/۳ به دست آمد. در سطح اطمینان ۹۰ درصد صحت فرضیه فوق را با فرض نرمال بودن عملکرد آزمون کنید.

مهارت ادراکی

مهارت عملی

$$n_1 = 14$$

$$\bar{x}_1 = 43$$

$$s_1 = 1.1$$

$$1 - \alpha = 0.9$$

$$\mu_1 > \mu_2$$

$$n_2 = 15$$

$$\bar{x}_2 = 40.7$$

$$s_2 = 1.3$$

توزیع نرمال و انحراف
معیارهای جامعه ها نامعلوم

$$\begin{cases} H_0: \mu_1 \leq \mu_2 \\ H_1: \mu_1 > \mu_2 \end{cases}$$

$$\begin{aligned} df &= n_1 + n_2 - 2 \\ &= 14 + 15 - 2 \\ &= 27 \end{aligned}$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

ادعای مسئله

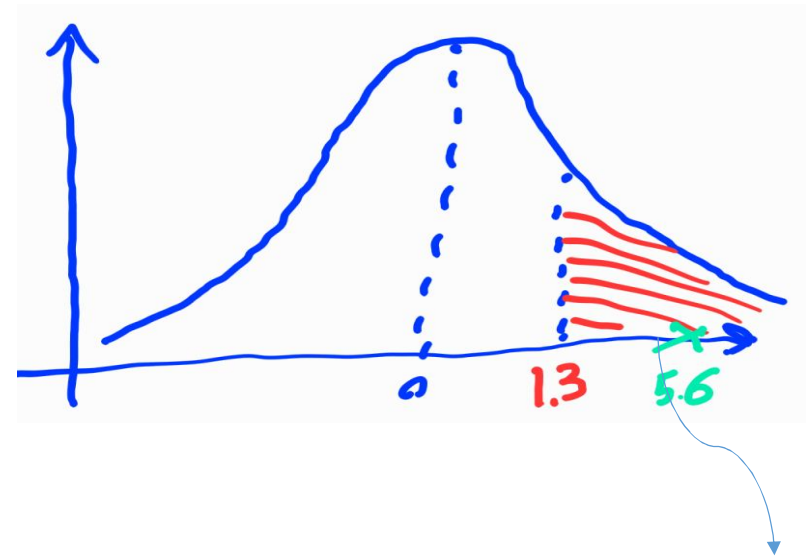
واحد رصفهان خورگان

$$= \frac{(43 - 40.7) - 0}{\sqrt{\frac{(1.1)^2}{14} + \frac{(1.3)^2}{15}}}$$

$$= 5.16$$

$$t_{0.01, (26)} = 1.31$$

$$df = \frac{\left(\frac{(1.1)^2}{14} + \frac{(1.3)^2}{15}\right)^2}{\frac{\left(\frac{(1.1)^2}{14}\right)^2}{14-1} + \frac{\left(\frac{(1.3)^2}{15}\right)^2}{15-1}} \approx 26$$



مقدار آماره آزمون در ناحیه بحرانی قرار می گیرد. بنابراین، در سطح خطای ده درصد فرض صفر رد می شود و ادعای مسئله در فرض یک تائید می شود. لذا می توان گفت عملکرد مدیران با مهارت ادراکی بالا بیشتر از عملکرد مدیران با مهارت عملی بالاست.



تمرین

در کارخانه ای دو خط تولید مستقل از هم، اشیاء یکسان تولید می کنند و وزن این اشیاء بر طبق قانون نرمال با پارامترهای نامعلوم توزیع می شود. نمونه ای به حجم های ۲۰۰ و ۴۰۰ از خط تولید اول و دوم انتخاب می شود. میانگین و واریانس در نمونه خط اول به ترتیب ۱۹۸ و ۲۰ و میانگین و واریانس در نمونه خط دوم به ترتیب ۲۰۰ و ۲۵ به دست آمده اند. با فرض یکسان بودن واریانسها در دو جامعه در سطح خطای ۰.۰۵ آزمون کنید که میانگینها در دو خط تولید یکسان می باشد.



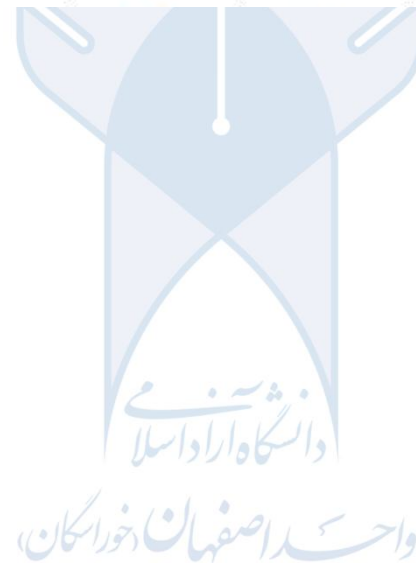


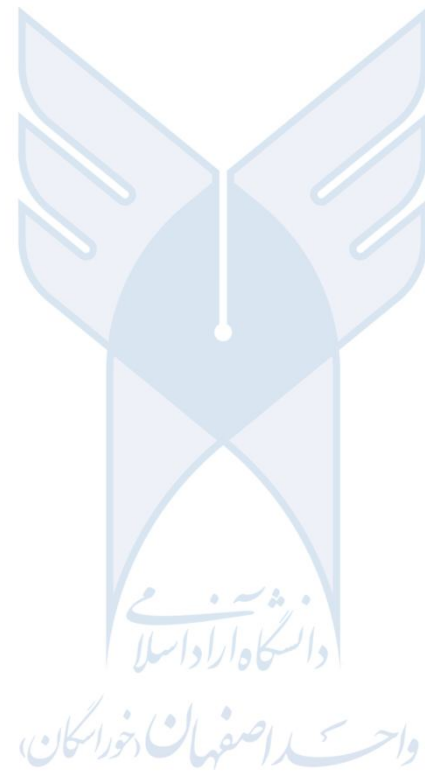
تمرین



برای مقایسه شادی دختران و پسران دانشگاهی به طور تصادفی از هر گروه 5 نفر انتخاب شده اند و میزان شادی اعضای نمونه با استفاده از یک پرسشنامه روانشناسی اندازه گیری شد نتایج به شرح زیر می باشد. با فاصله اطمینان 95 درصد آزمون نمایید که آیا شادی دختران و پسران با هم تفاوتی دارد یا نه؟ میزان شادی دارای توزیع نرمال و واریانس‌ها یکسان است.

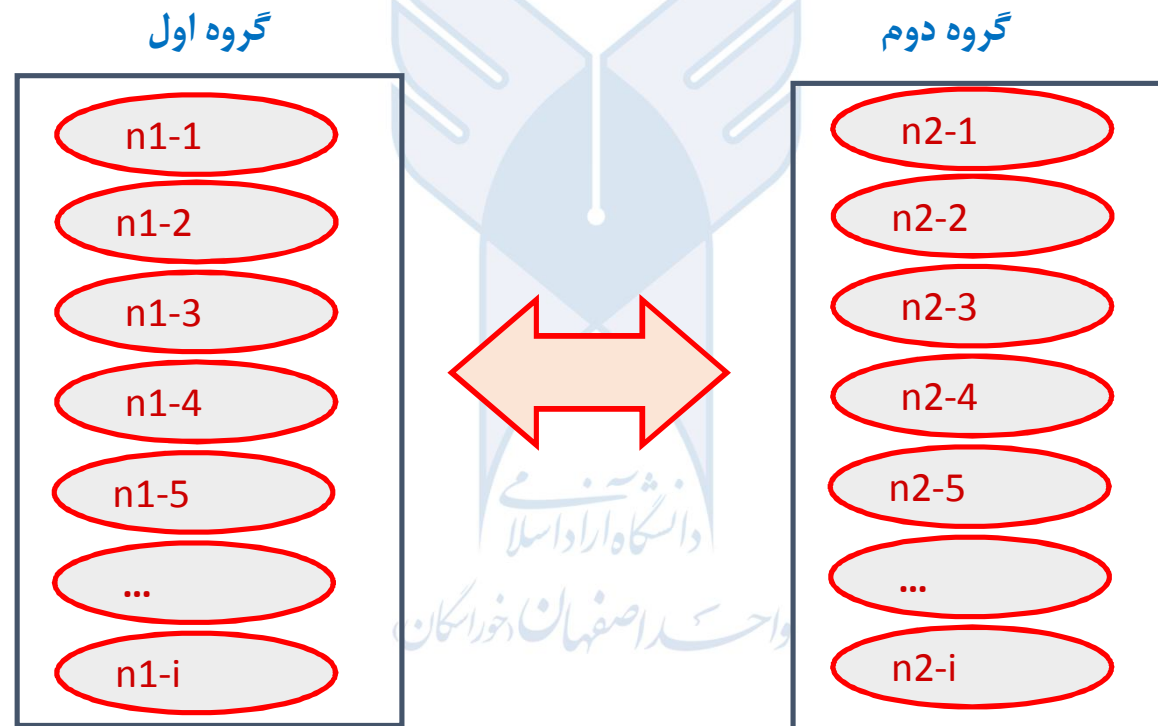
دختر	65	59	53	52	51
پسر	45	41	40	38	37





آزمون مقایسه زوجها (*paired sample test*)

استفاده از این آزمون زمانی است که محقق بخواهد اعضای دو مجموعه متناظر را که دارای تعداد نمونه یکسان هستند مورد مقایسه قرار دهد. به عنوان مثال یک مجموعه ای از شاخص ها را در قبل و بعد از یک اقدام مورد مقایسه قرار دهد. (نمونه‌های وابسته)



- مراحل عمومی این آزمون همانند آزمونهای قبلی می باشد با این تفاوت که ملاک آزمون از فرمول زیر بدست می آید.

$$t = \frac{\bar{d}}{s_{\bar{d}}}$$

$$\bar{d} = \frac{\sum d_i}{N}$$

$$s_{\bar{d}} = \frac{s_d}{\sqrt{n}}$$

$$s_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n-1}}$$

دانشگاه آزاد اسلامی
واحد رصفهان خوراگان

مثال

در راستای ایجاد وابستگی شغلی و بالا بردن انگیزه کاری در سازمان طرح جدیدی پیشنهاد شده است. ابتدا این طرح برای ۵ نیروی سازمانی اجرا شده و نتایج آن به شرح جدول زیر است. آیا می توان ادعا نمود اجرای این طرح سبب افزایش انگیزه کاری می شود؟ سطح معنی داری را ۵ درصد در نظر بگیرید.

شماره	① انگیزه کاری قبل از اجرای طرح	② انگیزه کاری بعد از اجرای طرح	d_i	$d_i - \bar{d}$	$(d_i - \bar{d})^2$
۱	۱۱۵	۱۲۷	$127 - 115 = 12$	$12 - 2 = 10$	۱۰۰
۲	۱۱۲	۱۱۵	$115 - 112 = 3$	$3 - 2 = 1$	۱
۳	۱۰۷	۱۰۶	$106 - 107 = -1$	$-1 - 2 = -3$	۹
۴	۱۱۹	۱۱۷	$117 - 119 = -2$	$-2 - 2 = -4$	۱۶
۵	۱۰۴	۱۰۲	$102 - 104 = -2$	$-2 - 2 = -4$	۱۶

$$\bar{d} = \frac{\sum d_i}{n} = \frac{10}{5} = 2$$

$$s_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n - 1}}$$

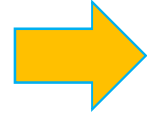
$$= \sqrt{\frac{142}{5 - 1}} = 5.95$$

$$\mu_d = \mu_2 - \mu_1$$

ادعای مسئله



$$\mu_2 > \mu_1$$



$$\mu_d = \mu_2 - \mu_1 > 0$$

$$\alpha = 0.05$$

$$df = n - 1 = 5 - 1 = 4$$

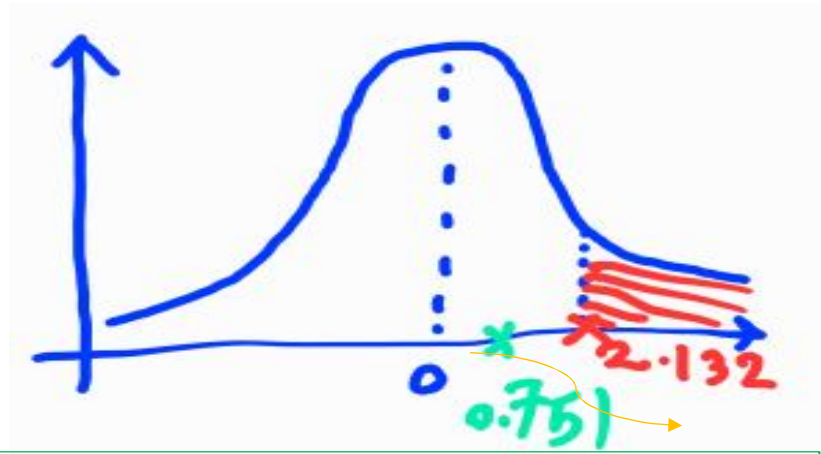
$$t_{0.05, (4)} = 2.132$$

$$\begin{cases} H_0: \mu_d \leq 0 \\ H_1: \mu_d > 0 \end{cases}$$

ادعای مسئله

$$s_d = \frac{s_d}{\sqrt{n}} = \frac{5.95}{\sqrt{5}} = 2.66$$

$$t = \frac{\bar{d}}{s_d} = \frac{2}{2.66} = 0.751$$



مقدار آماره آزمون در ناحیه پذیرش قرار می گیرد. بنابراین، در سطح خطای پنج درصد فرض صفر پذیرفته و ادعای مسئله در فرض یک تائید نمی شود. بنابراین اجرای طرح تاثیر گذار در افزایش انگیزه کاری نیست.

مثال

نمرات پیش آزمون و پس آزمون چهار فراگیر در یک دوره آموزشی به شکل زیر می باشد:

با فرض نرمال بودن نمرات، آیا فرضیه زیر در سطح اطمینان ۹۸ درصد مورد تایید است؟

“دوره آموزشی مورد نظر اثربخش است”

	۱	۲	۳	۴
پیش آزمون	۱۷	۸	۴	۱۵
پس آزمون	۱۵	۱۵	۴	۱۸
d_i	-۲	۷	۰	۳
$d_i - \bar{d}$	-۴	۵	-۲	۱
$(d_i - \bar{d})^2$	۱۶	۲۵	۴	۱

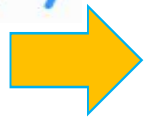
$$\bar{d} = \frac{\sum d_i}{n} = \frac{8}{4} = 2$$

$$S_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n - 1}}$$

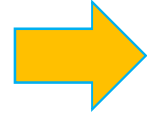
$$= \sqrt{\frac{46}{4 - 1}} = 3.91$$

$$\mu_d = \mu_2 - \mu_1$$

ادعای مسئله



$$\mu_2 > \mu_1$$



$$\mu_d = \mu_2 - \mu_1 > 0$$

$$\begin{cases} H_0: \mu_d \leq 0 \\ H_1: \mu_d > 0 \end{cases}$$

ادعای مسئله

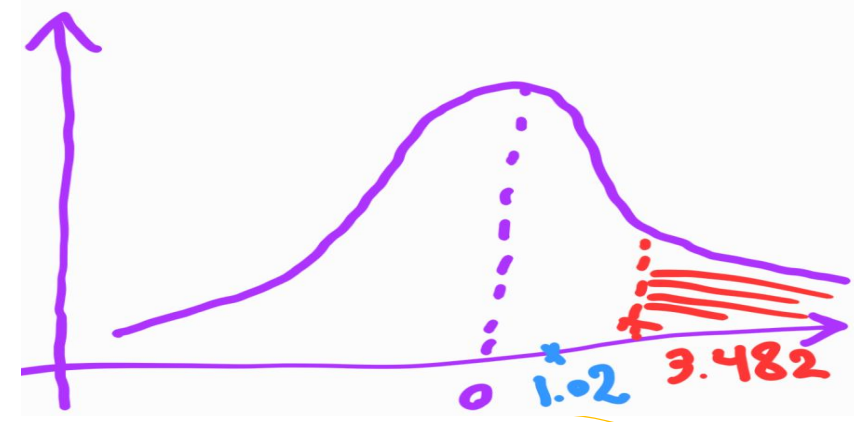
$$t = \frac{\bar{d}}{s_d}$$

$$= \frac{2}{\frac{3.91}{\sqrt{4}}} = 1.02$$

$$\alpha = 0.02$$

$$df = n - 1 = 4 - 1 = 3$$

$$t_{0.02, (3)} = 3.482$$



مقدار آماره آزمون در ناحیه پذیرش قرار می گیرد. بنابراین، در سطح خطای دو درصد فرض صفر پذیرفته و ادعای مسئله در فرض یک تائید نمی شود. پس دوره مورد نظر اثر بخش نیست.

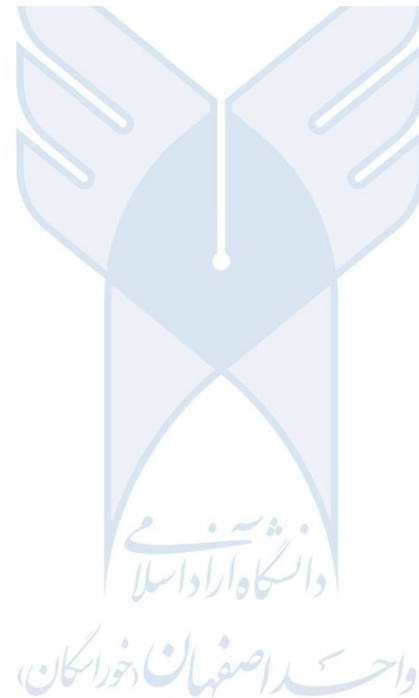
عنوان: تفاضل میانگین های دو جامعه

تمرین



محققى علاقمند است تاثير روش تدریس الف و روش تدریس ب را بر یادگیری درس آمار را با هم مقایسه کند از این رو یک گروه از آزمودنی را شش ماه با روش الف و شش ماه با روش ب آموزش می دهد پس از هر آموزش ، آزمون پیشرفت تحصیلی اجرا شده است نتایج در زیر آمده است آزمون را با اطمینان 95 درصد انجام دهید. توزیع نمرات نرمال است.

روش الف	روش ب
x	y
6	4
10	8
7	5
6	7
5	3
10	7



آزمون فرض مقایسه نسبتها در دو جامعه

استفاده از این آزمون زمانی است که فرضیه محقق در مورد نسبتهای دو جامعه باشد. یا به عبارت دیگر بخواهد نسبت شاخص مورد بررسی را در دو سازمان مورد مقایسه قرار دهد.

• مرحله اول) تعریف H_0 و H_1

حالت اول

$$\begin{cases} H_0 : p_1 = p_2 \\ H_1 : p_1 \neq p_2 \end{cases}$$

حالت دوم

$$\begin{cases} H_0 : p_1 \geq p_2 \\ H_1 : p_1 < p_2 \end{cases}$$

حالت سوم

$$\begin{cases} H_0 : p_1 \leq p_2 \\ H_1 : p_1 > p_2 \end{cases}$$

• مرحله دوم) محاسبه آماره آزمون یا ملاک آزمون

$$z = \frac{(\bar{p}_1 - \bar{p}_2)}{\sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1} + \frac{\bar{p}_2(1-\bar{p}_2)}{n_2}}}$$

و یا

$$z = \frac{(\bar{p}_1 - \bar{p}_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$\bar{p}_1 = \frac{X_1}{n_1}$$

$$\bar{p}_2 = \frac{X_2}{n_2}$$

$$\bar{p} = \frac{X_1 + X_2}{n_1 + n_2}$$

• مرحله سوم) تعیین ناحیه بحرانی

• مرحله چهارم) تصمیم گیری

مثال

به منظور مقایسه نسبت محصولات درجه یک در دو موسسه تولیدی، از موسسه اول ۴۰۰ واحد و از موسسه دوم ۱۰۰ واحد محصول انتخاب می شود و ملاحظه می شود که در موسسه اول ۲۸۰ محصول و در موسسه دوم ۶۰ واحد محصول درجه یک بوده اند. آزمون کنید که آیا محصولات درجه یک موسسه اول بیش از موسسه دوم است؟

$$n_1 = 400$$

$$n_2 = 100$$

$$x_1 = 280$$

$$x_2 = 60$$

$$\bar{p}_1 = \frac{x_1}{n_1} = \frac{280}{400} = 0.7$$

$$\bar{p}_2 = \frac{x_2}{n_2} = \frac{60}{100} = 0.6$$

$$Z = \frac{0.7 - 0.6}{\sqrt{\frac{0.7(0.3)}{400} + \frac{0.6(0.4)}{100}}} = 1.84$$

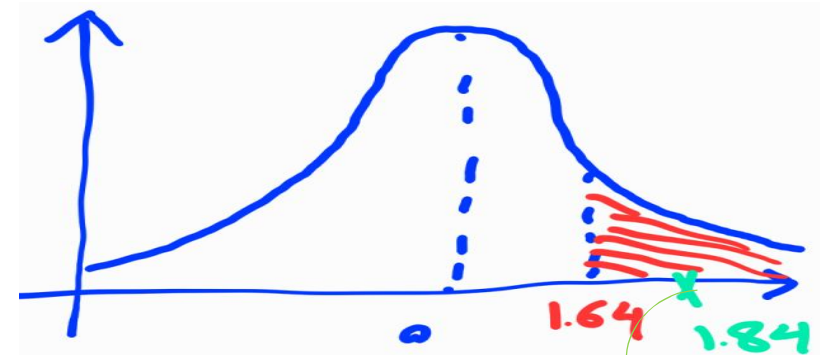
$\alpha = 0.05$

ادعای مسئله



$$P_1 > P_2$$

$$\begin{cases} H_0: P_1 \leq P_2 \\ H_1: P_1 > P_2 \end{cases}$$



مقدار آماره آزمون در ناحیه بحرانی قرار می گیرد. بنابراین، در سطح خطای پنج درصد فرض صفر رد می شود و ادعای مسئله در فرض یک تائید می شود.



تمرین

یک نمونه تصادفی ۱۰۰ تایی از مردان و ۱۰۰ تایی از زنان دانشکده انتخاب شده و از آنها در مورد داشتن ماشین شخصی سوال شده است. اگر ۳۱ نفر از مردان و ۲۴ نفر از زنان بگویند که صاحب ماشین شخصی هستند، می توانیم نتیجه بگیریم که مردان بیش از زنان ماشین شخصی دارند؟ سطح تشخیص را ۱ درصد در نظر بگیرید.





از توحه سما سپاسگزارم

دانشگاه آزاد اسلامی
واحد رصفهان خوراگان