

①

حالت سوم: توزیع نامعلوم (غیر نرمال)، واریانس نامعلوم:

در صورتی که بخواهیم نسبت به مقایسه میانگین های جامعه با اختلاف میانگین ها با فرض اولاً: مجهول (غیر نرمال) بودن متغیرها در جامعه ثانیاً: معلوم بودن واریانس های جامعه ضاوت کنیم:

(۱) نمونه های تصادفی به اندازه کافی بزرگ انتخاب می کنیم $n_1, n_2 \geq 30$ و $n_1, n_2 \rightarrow +\infty$

(۲) $\bar{X}_1$ و $\bar{X}_2$ میانگین های نمونه ای را محاسبه می کنیم

(۳) ملاک آزمون $\Rightarrow T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim N(0,1)$

(۴) مرتبه آزادی کوانیل: $n_1 + n_2 - 2$

مثال ۱:

فرض کنید توزیع دفعات تأخیر کارکنان زن و مرد مجهول (غیر نرمال) و در یک نمونه تصادفی از کارکنان زن و مرد به حجم ۱۰۰ نفر دفعات تأخیر به صورت زیر توزیع شده باشد:

جمع	۰	۱	۲	۳	۴	جمع
تعداد کارکنان	۱۰	۲۵	۳۵	۲۰	۱۰	۱۰۰

$$d_1^2 = 4 \quad d_2^2 = 9$$

جمع	۲	۳	۴	۵	۶	جمع
تعداد کارکنان	۱۰	۲۰	۴۰	۲۰	۱۰	۱۰۰

مطلوب است در سطح ۵ درصد خطای نوع اول:

آزمون ادعای منبسط تر بودن کارکنان زن

حل:

ابتدا میانگین های نمونه ای را محاسبه می کنیم:

x_1	n_1	$x_1 n_1$
۰	۱۰	۰
۱	۲۵	۲۵
۲	۳۵	۷۰
۳	۲۰	۶۰
۴	۱۰	۴۰
جمع	۱۰۰	۱۹۵

$$\bar{X}_1 = \frac{195}{100} = 1.95$$

$$\bar{X}_2 = \frac{400}{100} = 4$$

x_2	n_2	$x_2 n_2$
۲	۱۰	۲۰
۳	۲۰	۶۰
۴	۴۰	۱۶۰
۵	۲۰	۱۰۰
۶	۱۰	۶۰
جمع	۱۰۰	۴۰۰

⑤ $\mu_1 < \mu_2$

1) $\begin{cases} H_0: \mu_1 - \mu_2 \geq 0 \\ H_1: \mu_1 - \mu_2 < 0 \end{cases}$

2) $\begin{cases} \text{توزیع دیتا تا آخر جدول} \\ d_1^2 = 4, d_2^2 = 9 \\ n_1, n_2 \rightarrow \infty, \alpha = 5\% \end{cases}$

3) $T = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim N(0,1)$



5) $T = \frac{(11.95 - 4) - 0}{\sqrt{\frac{4}{11} + \frac{9}{11}}} = \frac{-7.95}{1.3406} \approx -5.93$

4) $T = -5.93 < -1.96 \Rightarrow R_{H_0} \text{ و } A_{H_1}$

نتیجه گیری:

به احتمال ۹۵ درصد کارخان زن منضبط ترند. به احتمال ۹۵ درصد متوسط دفعات تاخیر کارخان زن کمتر است.

حالت چهارم: هر دو مجهول

اگر میخواهیم نسبت به مقایسه میانگین های جامعه یا اختلاف میانگین ها قضاوت کنیم درحالی که:
اولاً: توزیع متغیر در هر ۲ جامعه مجهول (غیرنرمال)
ب) نمونه های تصادفی به اندازه کافی بزرگ انتخاب می کنیم
۱۲ $\bar{X}_1, \bar{X}_2$ میانگین های نمونه و s_1^2, s_2^2 واریانس های نمونه ای را محاسبه می کنیم

$n_1, n_2 \geq 30, n_1, n_2 \rightarrow +\infty$

$(\bar{X}_1 - \bar{X}_2) \sim N(\mu_1 - \mu_2, \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}) \Rightarrow T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim N(0,1)$
 $n_1, n_2 \rightarrow +\infty$

$n_1, n_2 \rightarrow +\infty$

x_1	n_1	$x_1 n_1$	$x_1^2 n_1$
0	11	0	0
1	25	25	25
2	25	50	100
3	20	60	180
4	10	40	160
Σ	111	195	505

x_2	n_2	$x_2 n_2$	$x_2^2 n_2$
2	10	20	40
3	20	60	180
4	40	160	640
5	20	100	500
6	10	60	360
Σ	110	400	1720

حل:
 $\bar{X}_1, \bar{X}_2$
 s_1^2, s_2^2
را محاسبه می کنیم.

مثال ۲:

در مثال قبل فرض کنید واریانس ها مجهول است، مجدداً ادعای مطمح شده رازمون کنید.

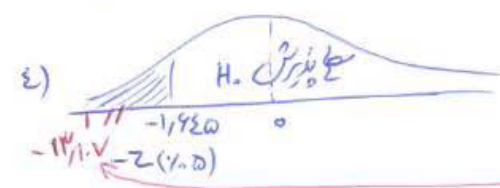
$\bar{X}_1 = \frac{195}{111} = 1.756, s_1^2 = \frac{505}{111} - 1.756^2 = 1.2475$
 $\bar{X}_2 = \frac{400}{110} = 3.636, s_2^2 = \frac{1720}{110} - 3.636^2 = 1.12$

3) $\mu_1 < \mu_2$

1) $\begin{cases} H_0: \mu_1 - \mu_2 \geq 0 \\ H_1: \mu_1 - \mu_2 < 0 \end{cases}$

2) $\begin{cases} \text{توزیع دفعات تأخیر محمول غیر نرمال است} \\ \text{طریقی که محمول است} \end{cases}$
 $\alpha = 5\%$
 $n_1, n_2 \rightarrow +\infty$

3) $T = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim N(0,1)$
 $n_1, n_2 \rightarrow +\infty$



4) $T = \frac{(1.95 - 4) - 0}{\sqrt{\frac{1.245^2}{12} + \frac{1.2^2}{12}}} = \frac{-2.05}{\sqrt{.1064}} \approx -13/10.7$

5) $0, T = -13/10.7 < -1.945 \Rightarrow R_{H_0}, A_{H_1}$

17) نتیجه گیری:
 - ۹۵ درصد کارکنان نرخ منقبض ترند.
 - ۹۵ درصد متوسط دفعات تأخیر کارکنان نرخ کمتر است.

تمرین ۱:

اگر توزیع بازده لام شرکت A و B محمول یا غیر نرمال باشد و طریق ۱ روز به تصادف به صورت زیر توزیع شده باشد

شرکت A	
بازده لام به درصد	تعداد روز
۱۰ - ۱۲	۱۰
۱۲ - ۱۴	۲۰
۱۴ - ۱۶	۴۰
۱۶ - ۱۸	۲۰
۱۸ - ۲۰	۱۰
جمع	۱۰۰

شرکت B	
بازده لام به درصد	تعداد روز
۸ - ۱۲	۱۰
۱۲ - ۱۶	۲۰
۱۶ - ۲۰	۴۰
۲۰ - ۲۴	۲۰
۲۴ - ۲۸	۱۰
جمع	۱۰۰

در هر یک از حالات متقل زیر:
 الف) واریانس با معلوم $k_1^2 = 1, k_2^2 = 4$
 ب) واریانس ها محمول مرابسه.

در سطح ادرسه خطای نوع اول ادعای زیر را آزمون کنید:

الف) متوسط بازده شرکت A کمتر است، (شرکت A بازدهی کمتر دارد)
 ب) بازدهی بهام ۲ شرکت به طور متوسط متفاوت اند.

(4)

پارامتر دوم: مقایسه نسبت ها

اگر بخواهیم به عنوان تست پارامترهای از جامعه که دارای ویژگی خاصی هستند مقایسه کنیم؛

$$P_1 < P_2 \Rightarrow P_1 - P_2 < 0 \quad \text{و} \quad \dots$$

$$n_1, n_2 \geq 30 \Rightarrow n_1, n_2 \rightarrow +\infty$$

$$\hat{P}_1 = \frac{x_1}{n_1} \quad \text{و} \quad \hat{P}_2 = \frac{x_2}{n_2}$$

$$(\hat{P}_1 - \hat{P}_2) \sim N(P_1 - P_2, \sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1} + \frac{\hat{P}_2(1-\hat{P}_2)}{n_2}})$$

$$\Rightarrow T = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1} + \frac{\hat{P}_2(1-\hat{P}_2)}{n_2}}} \xrightarrow{n_1, n_2 \rightarrow +\infty} (1)$$

(5) مزبور از جدول کوئینل های توزیع نرمال است که در دست می آید.

مثال 3:

اگر جهت ارزیابی بگه فرصت اشتغال به منطقه جغرافیایی سکونت در منطقه شمال و جنوب تهران به ترتیب ۱۰۰۰ و ۱۰۰۰۰ نفر از بین افراد واجد شرایط اشتغال انتخاب (از آنها ۹۰۰ و ۸۵۰۰ نفر شاغل باشند) مطلوب است در سطح درصد خطای نوع اول آزمون ادعا های؛
الف) بگه نداری فرصت اشتغال به منطقه سکونت (تایید کردن نسبت اشتغال)
ب) بهتر بودن سطح اشتغال در شمال تهران

حل!

ابتدا فرارهای بگه را محاسبه می کنیم:

$$\hat{P}_1 = \frac{900}{10000} = 0.09$$

$$\hat{P}_2 = \frac{8500}{10000} = 0.85$$

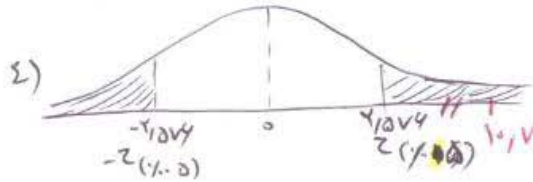
۵)

$$1) \begin{cases} H_0: P_1 - P_2 = 0 \\ H_1: P_1 - P_2 \neq 0 \end{cases}$$

$$2) \begin{cases} n_1, n_2 \rightarrow +\infty \\ \alpha = 1\% \quad \alpha/2 = 0.5\% \\ \hat{P}_1 = 1/9 \quad \hat{P}_2 = 1/15 \end{cases}$$

$$3) T = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1} + \frac{\hat{P}_2(1-\hat{P}_2)}{n_2}}} \sim N(0, 1)$$

فرض الف) $P_1 = P_2$



$$5) T = \frac{(1/9 - 1/15) - 0}{\sqrt{\frac{1/9 * 8/9}{11} + \frac{1/15 * 14/15}{11}}} = \frac{1.05}{2.0444} = 1.073$$

$$6) T = 1.073 > 2.1574 \Rightarrow R_{H_0}, A_{H_1}$$

۱۷ نتیجه گیری

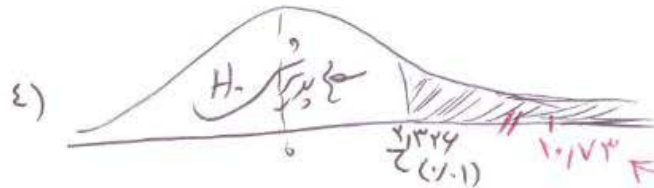
نتیجه گیری: در صورت استغال نسبت استغال در منطقه برابر است. به نظر می آید که نسبت در منطقه

$$1) \begin{cases} H_0: P_1 - P_2 \leq 0 \\ H_1: P_1 - P_2 > 0 \end{cases}$$

$$2) \begin{cases} n_1, n_2 \rightarrow +\infty \\ \alpha = 1\% \\ \hat{P}_1 = 1/9 \quad \hat{P}_2 = 1/15 \end{cases}$$

$$3) T = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1} + \frac{\hat{P}_2(1-\hat{P}_2)}{n_2}}} \sim N(0, 1) \quad n_1, n_2 \rightarrow +\infty$$

فرض ب) $P_1 > P_2$



$$5) T = \dots = 1.073$$

$$6) T = 1.073 < 2.1574 \Rightarrow R_{H_0}, A_{H_1}$$

۱۷ نتیجه گیری

نتیجه گیری: در صورت استغال در منطقه استغال از حد بزرگتر است. نسبت استغال منطقه استغال بزرگتر است.

نکته:

۱) می توان مصاحبه را بر حسب درصد بیان کرد

تمرین ۲:

در مثال قبل

- ۱) ادعاها را بر حسب درصد استغال مجدداً آزمون کرده نتایج را مقایسه کنید.
- ۲) ادعاها را بر حسب نسبت بیان مجدداً آزمون کرده و مقایسه کنید.

۲) می توان محاسبه آزمون را بر حسب نسبت بیان کرد