آزمون های آماری ۵

الف- آزمایی‌ها را در اختیار داشته باشیم.

(واریانس مجهول)

توزیع جمعیت نرمال

$$\begin{cases} H_0: \mu_x = \mu_0 \\ H_1: \mu_x \neq \mu_0 \end{cases} \quad (1) \qquad \begin{cases} H_0: \mu_x \leq \mu_0 \\ H_1: \mu_x > \mu_0 \end{cases} \quad (2) \qquad \begin{cases} H_0: \mu_x \geq \mu_0 \\ H_1: \mu_x < \mu_0 \end{cases} \quad (3)$$

$$T = \frac{\bar{X} - \mu_0}{\sqrt{\frac{S_n^2}{n}}} \overset{H_0}{\sim} t(n-1) \qquad (1) \quad \text{if } |T_0| \geq t_{1-\alpha/2}(n-1) \Rightarrow RH_0$$

$$\text{if } p\text{-value} = P(|T| \geq |T_0|) = 2P(T \geq |T_0|) < \alpha \Rightarrow RH_0$$

$$(2) \quad \text{if } T_0 \geq t_{1-\alpha}(n-1) \Rightarrow RH_0$$

$$\text{if } p\text{-value} = P(T \geq T_0) < \alpha \Rightarrow RH_0$$

$$(3) \quad \text{if } T_0 \leq -t_{1-\alpha}(n-1) \Rightarrow RH_0$$

$$\text{if } p\text{-value} = P(T \leq T_0) < \alpha \Rightarrow RH_0$$

در بسیاری از نرم افزارها مانند SAS نتیج آزمون شماره (1) را انجام می دهند، p-value مربوط به آن را گزارش می کنند. برای بررسی آزمون های (2) و (3) در SAS به طریق زیر عمل می کنیم:

- آر p-value مربوط به آزمون (i) را با PV_i نشان دهیم آنگاه:

$PV_1 = 2P(T \geq |T_0|)$ ← از خروجی در بسیاری از نرم افزارها به دست می آید

$$PV_2 = \begin{cases} \frac{PV_1}{2} & \text{if } T_0 \geq 0 \\ 1 - \frac{PV_1}{2} & \text{if } T_0 < 0 \end{cases}, \quad PV_3 = \begin{cases} 1 - \frac{PV_1}{2} & \text{if } T_0 \geq 0 \\ \frac{PV_1}{2} & \text{if } T_0 < 0 \end{cases}$$

برهان:

$$PV_1 = 2P(T \geq |T_0|), \text{ if } T_0 \geq 0 \Rightarrow PV_1 = 2P(T \geq T_0)$$

$$\Rightarrow \underbrace{P(T \geq T_0)}_{PV_2} = \frac{PV_1}{2}$$

$$\text{if } T_0 \leq 0 \Rightarrow PV_1 = 2P(T \geq -T_0) = 2(1 - P(T \geq t_0))$$

$$\Rightarrow \underbrace{P(T \geq t_0)}_{PV_2} = 1 - \frac{PV_1}{2}$$

در مورد PV_3 نیز به همین شیوه می توان استنباط کرد. (این مطلب جواب تمرین اینجا را دارد.)

از دستور means یا univariate می‌توان برای بدست آوردن PV1 استفاده کرد.

```
data a;
input X;
cards;
11
12
13
14
18
12
19
;
Proc univariate mu0=15;
Var x;
run;
```

آزمون شود $\mu_0 = 0$ در نظر گرفته می‌شود.

```
data a;
input X;
Y = X-15;
cards;
11
12
13
14
18
12
19
;
Proc means T PRT noprint;
Var Y;
run;
```

برای بدست آوردن PV2 یا PV3 - طریق زیر عمل می‌کنیم:

```
Proc univariate T PRT mu0=15 noprint;
Var x;
output out=data_PV1 PRT=PV1 T=T0;
run;
```

در دستور univariate هم مانند means با قسمت output خروجی را ذخیره می‌کنیم و ... تا شود در این جا Var y; را به کار می‌بریم.

```
data b;
set data_PV1;
if T0>=0 then PV2 = PV1/2;
else PV2 = 1 - PV1/2;
run;
Proc Print data=b;
Var PV2;
run;
```

* از دستور TTest نیز می‌توان برای انجام آزمون فوق استفاده کرد:

```
Proc TTest H0=15;
Var x;
run;
```

ب- آزمون جمعیت دانشجویی

(واریانس‌ها نامعلوم و جمعیت‌ها توزیع نرمال دارند.)

۱- آزمون‌های مورد بررسی در دو جمعیت زوجی باشند:

$$\begin{cases} H_0: \mu_x - \mu_y = \delta \\ H_1: \neq \delta \end{cases} \quad (1) \qquad \begin{cases} H_0: \mu_x - \mu_y \leq \delta \\ H_1: > \delta \end{cases} \quad (2) \qquad \begin{cases} H_0: \mu_x - \mu_y \geq \delta \\ H_1: < \delta \end{cases} \quad (3)$$

روش حل:

$$D = X - Y \qquad \mu_d = \mu_x - \mu_y$$

$$\begin{cases} H_0: \mu_d = \delta \\ H_1: \neq \end{cases} \quad (1) \qquad \begin{cases} H_0: \mu_d \leq \delta \\ H_1: > \delta \end{cases} \qquad \begin{cases} H_0: \mu_d \geq \delta \\ H_1: < \delta \end{cases} \quad (3)$$

$$T = \frac{\bar{D} - \delta}{\sqrt{\frac{S_D^2}{n}}} \overset{H_0}{\sim} t(n-1)$$

استباط به حالت یک جمعیت می‌شود.

به دو طریق می‌توان آزمون‌های فوق را انجام داد:

```
data c;
input x y;
D = X - Y;
cards;
11  12
18  12
14  16
15  18
19  20
18  19
;
proc univariate mu0 = 2;
var D;
run;
```

بجای این دو دستور می‌توان از دستور means از proc TTests استفاده کرد.

```
data c;
input X Y;
cards;
11 12
18 12
14 16
15 18
19 20
18 19
;
proc TTest H0=2;
Paired X*Y;
run;
```

* تفاوت روش‌های فوق در این است که با دستور univariate می‌توان دستور output به کار برد و p-value آزمون‌های یک طرفه را آزمون داد ولی با دستور TTest، باید p-value را به صورت دستی حساب کرد.

* به لگاه هر دو روش می‌توان فاصله اطمینان برای $\mu_d = \mu_x - \mu_y$ به دست آورد. در این حالت با توجه به آزمون alpha در دستورهای می‌توان ضریب اطمینان دلخواه را تعیین کرد.

۲- آزمون مقایسه‌ی میانگین دو جامعه‌ی مستقل با فرض نرمال بودن

الف- آزمون $\sigma_x^2 = \sigma_y^2$ روش pooled

$$T = \frac{\bar{X} - \bar{Y} - \delta}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \overset{H_0}{\sim} t(n_1+n_2-2)$$

$$S_p^2 = \frac{(n_1-1)S_x^2 + (n_2-1)S_y^2}{n_1+n_2-2}$$ pooled

(1) if $|T| \geq t_{1-\frac{\alpha}{2}}(n_1+n_2-2) \Rightarrow RH_0$.

if p-value $= P(|T| \geq |T_0|) = 2P(T \geq |T_0|) \leq \alpha \Rightarrow RH_0$.

(2) if $T \geq t_{1-\alpha}(n_1+n_2-2) \Rightarrow RH_0$.

if p-value $= P(T \geq T_0) \leq \alpha \Rightarrow RH_0$.

(3) if $T \leq -t_{1-\alpha}(n_1+n_2-2) \Rightarrow RH_0$.

if p-value $= P(T \leq T_0) \leq \alpha \Rightarrow RH_0$.

برای انجام آزمون‌های فوق مربوط به دو جامعه‌ی مستقل از نرم‌افزار SAS استفاده می‌کنیم. یکی از طریق وارد کردن داده‌ها به برنامه‌ست زیر با شرط:

```
data d;
input X Code;
cards;
11 1
18 1
14 1
15 1
19 1
18 1
12 2
12 2
16 2
18 2
20 2
19 2
;
run;
```

```
proc TTEST H0=2;
class code;
var X;
run;
```

* در رویه‌ی TTEST گزینه‌های by، freq، weight را نیز می‌پذیرد.

* این رویه برای پارامترهای μ_x، μ_y، $\mu_x - \mu_y$ آزمون‌های زیر را انجام می‌دهد:

آزمون‌های یک‌طرفه و دوطرفه:

روش pooled: T, p-value

روش Satterthwaite: T, p-value

$$\begin{cases} H_0: \mu_x - \mu_y = \delta \\ H_1: \neq \delta \end{cases}$$

همچنین آزمون‌های دوطرفه‌ی واریانس‌ها:

T p-value

$$\begin{cases} H_0: \sigma_x^2 = \sigma_y^2 \\ H_1: \neq \end{cases}$$

براساس نتایج حاصل از آزمون برابری واریانس، چنانچه این فرض رد شود pooled یا روش Satterthwaite را استفاده کرد.

اگر $\sigma^2_x \neq \sigma^2_y$ باشد $T = \dfrac{\bar{X} - \bar{Y} - \delta}{\sqrt{\frac{S_x^2}{n_1} + \frac{S_y^2}{n_2}}} \overset{H_0}{\sim} t(f)$ Satterthwaite و...

$f = ?$

(1) if $|T| \geq \underset{1-\alpha/2}{t}(f) \Rightarrow RH.$

if $p\text{-value} = P(|T| \geq |T_0|) \leq \alpha \Rightarrow RH.$

(2) if $T \geq \underset{1-\alpha}{t}(f) \Rightarrow RH.$

if $p\text{-value} = P(T \geq T_0) \leq \alpha \Rightarrow RH.$

(3) if $T \leq -\underset{1-\alpha}{t}(f) \Rightarrow RH.$

if $p\text{-value} = P(T \leq T_0) \leq \alpha \Rightarrow RH.$