برنامه های جلسه سوم

داده های نمونه:

```
Data a;
```

trt	genotyp	x	y
a	1	12	2.5
a	2	10	3.6
a	3	14	5.7
a	4	19	6
a	5	22	8
b	1	11	3
b	2	8	4
b	3	13	5.2
b	4	14	6.2
b	5	20	7

Proc gchart:

این رویه نمودارهای فراوانی را رسم می کند. دستورات vbar و hbar نمودارهای فراوانی هر یک از متغیرهای مربوط به خود را به ترتیب عمودی و افقی رسم می کنند. البته در hbar علاوه بر رسم افقی فراوانی ساده، تجمعی، نسبی درصدی و تجمعی درصدی هم داده می شود. در این رویه می توان برای هر متغیر نوع نمودار فراوانی را مشخص کرد. در حالت عدم تعریف نوع نمودار تنها نمودار فراوانی ساده رسم می شود ولی می توان از گزاره های متفاوت با زدن یک اسلش (/) بعد از متغیر استفاده کرد مثلا در مثال زیر گفته شده برای متغیر y نمودار فراوانی ساده هم رسم شود و هم نمایش داده شود (اگر freq زده نمی شد فقط رسم می کرد) و برای متغیر x نمودار فراوانی ساده در دو طبقه فراوانی رسم شود. برای حدس تعداد طبقات در متغیر کمی پیوسته بایستی از فرمول $n = 2.5 \times N^{\frac{1}{4}}$ استفاده کرد که n تعداد طبقات و N تعداد کل افراد است. و برای تعیین فاصله طبقات بایستی از فرمول $i = \frac{R}{n}$ استفاده کرد که R دامنه و n تعداد طبقات است. در واقع n همان levels است. توجه شود فقط از یک دستور استفاده شود تا sas متوجه شود کدام دستور را انجام دهد. موارد دیگر شامل دستورات زیر است:

Percent: رسم نمودار فراوانی نسبی درصدی (درصدی)

Cfreq: رسم نمودار فراوانی تجمعی

Cpercent: رسم نمودار فراوانی تجمعی درصدی

Sum: رسم نمودار مجموع

Mean: رسم نمودار میانگین

Midpoints=values: نقاط میانی طبقات (میانه) را برای رسم نمودار فراوانی می دهد مثلا اعداد فلان و فلان نقاط میانی برای هر ستون فراوانی می شوند. برای بدست آوردن میانه در متغیر کمی پیوسته و گسسته و داده خام فرمول های متفاوتی موجود است که از آن فرمولها استفاده و میانه رو مشخص می کنیم.

Group=variable: این دستور گروهبندی را تعریف می کند (همانند by) و مثلا اگر group=trt شود نمودار فراوانی برای هر سطح تیمار رسم می شود.

```
data;
input trt$ genotyp x y;
cards;
;
proc gchart;
vbar y/freq;
hbar x/levels=2;
run;
```

Proc plot:

این رویه پراکنش یک متغیر را در برابر یک متغیر دیگر رسم می کند. در اینجا y روی x رسم شده است یعنی y محور عمودی و x محور افقی. اگر به جای رویه plot رویه gplot نوشته شود پلات های رسم شده حالت گرافیکی خواهند داشت.

```
data;
input trt$ genotyp x y;
cards;
;
proc plot;
plot y*x;
run;
```

داده های نمونه t-test:

A	B
34	45
22	38
15	52
27	48
37	25
41	39
24	51
19	46
26	55
36	46

رویه های t-test:

مقایسه میانگین یک جامعه با ادعای مطرح شده (**One sample t-test**):

در این دستور ابتدا فرض H0 را تعریف میکنیم و برنامه را ران می کنیم. در ttest همیشه حدود هم داده میشود. مثلا در شکل زیر، ستون اول میانگین اعداد را زده دو ستون بعدی حدود اطمینان به ترتیب پایین و بالای میانگین

رو زده و ستون بعد انحراف معیار را زده و دو ستون بعدی حدود اطمینان انحراف معیار همگی در سطح اطمینان ۹۵ درصد.

نکته: در آزمونهای t در SAS مهم نیست فرض H1 چیست و با رد مساوی نبودن آن می توان در مورد H0 اظهار نظر کرد.

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
8.0000	5.1201	10.8799	5.6013	4.1717	8.5248

```
data;
input A;
H0=X1-30;
Datalines;

;
proc ttest;
var H0;
run;
```

مقایسه میانگین دو جامعه جفت شده (paired t-test):

در حالت جفت شده از دو برای انجام t-test می توان از دو رویه استفاده کرد:

رویه زیر همان رویه means است. سپس متغیر d حاصل از تفاضل داده های متناظر تعریف می شود و مقادیر میانگین، مقدار t و مقدار احتمال برای آزمون فرض به ترتیب با mean، t و prt برای رویه تعریف می شود:

```
data;
input A B;
d=A-B;
cards;

;
proc means t prt mean stderr;
var d;
run;
```

رویه دیگر همان رویه t-test جفت شده است.

```
Data;
Input A B;
Datalines;

;
proc ttest;
paired A*B;
run;
```

مقایسه میانگین دو جامعه جفت نشده (two sample t-test):

برای انجام این نوع t-test بایستی داده ها به فرمت کلاسه بندی وارد شوند که در مثال زیر به شکل چند ردیفه نیز وارد کرده ایم:

group	y	group	y
A	34	B	45
A	22	B	38
A	15	B	52
A	27	B	48
A	37	B	25
A	41	B	39
A	24	B	51
A	19	B	46
A	26	B	55
A	36	B	46

```
data;
input group$ y@@;
Datalines;

;
proc ttest;
class group;
var y;
run;
```