

برنامه های جلسه دوم:
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 |

Data b;

| rep | trt | y1 | y2 |
|-----|-----|----|-----|
| 1 | A | 4 | 7.2 |
| 1 | B | 3 | 6.3 |
| 1 | C | 4 | 4.5 |
| 1 | D | 5 | 3.7 |
| 2 | A | 7 | 5.5 |
| 2 | B | 2 | 3.5 |
| 2 | C | 6 | 4.6 |
| 2 | D | 3 | 7.2 |
| 3 | A | 4 | 6.3 |
| 3 | B | 6 | 5.2 |
| 3 | C | 2 | 4.1 |
| 3 | D | 4 | 5.9 |

Proc print و انواع داده خوانی:

```
data a;  
input trt$ genotyp x y;  
datalines;  
  
;  
run;  
data b;  
input rep trt$ y1 y2;  
cards;  
;  
run;  
proc print;  
run;  
proc print data=a;  
run;
```

```
data a;  
input trt$ genotyp x y;  
datalines;  
;  
proc sort;  
by genotyp;  
run;
```

:Proc sort

:Proc means

در این پروگ انواع آماره های توصیفی قابل تعریف است.

این پروگ هم می تواند با دستور `class` و هم بدون آن تعریف گردد. در صورتی که `class` تعریف گردد یعنی میانگین گیری و برآورد سایر آماره ها بر مبنای متغیر موجود در `class` دسته بندی می شود و بر حسب سطوح این متغیر میانگین گیری می شود. مثلا عامل تیمار (`trt`) دارای دو سطح `a` و `b` است و اگر در `class` تیمار به عنوان فاکتور طبقه بندی تعریف شود داده ها به دو کلاس مذکور (سطوح) تقسیم (کلاس بندی) شده و بر حسب این دو سطح داده ها میانگین گیری می شود. در واقع روی ژنوتیپها در هر تیمار میانگین گیری می شود و دو میانگین بدست می آید. حال اگر `class` شامل هم تیمار و هم ژنوتیپ باشد میانگین های اثر متقابل تیمار در ژنوتیپ بدست می آید. البته در اینجا تکرار نداریم و اگر این کار را کنیم دیگر چیزی نیست که رویش میانگین گرفته شود و خود اعداد را می دهد اما اگر تکرار داشتیم روی تکرارها میانگین گیری می شد. این نکته نشان می دهد می توان از این راه انواع میانگین های اثرات متقابل را بدست آورد (راه دیگری به جز `lsmeans`). در ضمن در صورتی که به جای گزاره `class` از `by` استفاده شود نتیجه یکسان است. `Median`، `mode`، `range`، `uss` و `css` به ترتیب میانه، مد، دامنه تغییرات، `ss` تصحیح نشده و `ss` تصحیح شده را بدست می دهد. همچنین `maxdec` تعداد رقم اعشار برای چاپ عدد را تعریف می کند (در مثال زیر اعداد تا دو رقم اعشار چاپ می شوند).

```
data;
input trt$ genotyp x y;
cards;

;
proc means maxdec=2 mean std stderr min max range median mode sum var uss css
cv skewness kurtosis;
class trt;
var y;
run;
```

:proc univariate

این پروگ (رویه) علاوه بر گزارش برخی از آماره های توصیفی انواع پارامترهای تست نرمالیتی را نیز بدست می دهد. شاید کاربرد اصلی آن ۱. یافتن داده پرت و ۲. پی بردن به توزیع داده ها باشد.

این رویه می تواند همراه با `by` یا بدون آن باشد. اگر همراه با `by` باشد، `by` بعد از `var` تعریف می شود مثلا `trt by` باعث می گردد این رویه برای هر سطح تیمار جداگانه اجرا شود و نرمالیتی و سایر آماره ها را در هر سطح تیمار جداگانه بررسی کند.

```
data;
input trt$ genotyp x y;
cards;

;
proc univariate normal plot;
var x y;
```

```
histogram x y/normal(color=red w=5);  
run;
```

:Proc freq

این رویه جدول فراوانی را بدست می دهد. دقت شود این رویه فقط برای متغیرهای کمی گسسته کاربرد دارد و برای متغیرهای کمی پیوسته بهتر است از رویه **gchart** استفاده شود که مفصلا در جلسه بعد توضیح داده می شود.

```
data;  
input trt$ genotype x y;  
cards;  
;  
proc freq;  
table x y;  
run;
```