Ch3 Lab Results - Correlation & Regression

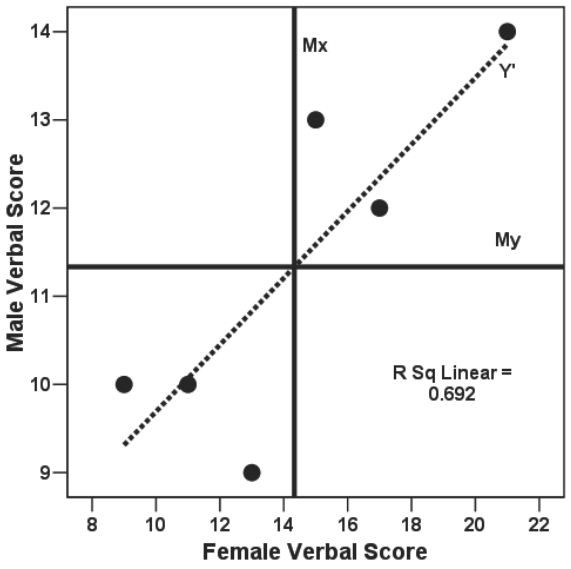
```
*1&2-enter and graph.
DATA LIST FREE/ fverb mverb.
BEGIN DATA
15 13   11 10   17 12    9 10   13  9   21 14
END DATA.

*Do with menus; Chart editor to add lines, text.

GRAPH /SCATT(BIVAR) = fverb WITH mverb.

*3-SPSS computations for correlation/regression.
COMPUTE fdev = fverb - 14.3333.
COMPUTE mdev = mverb - 11.3333.
COMPUTE fdev2 = fdev**2.
COMPUTE mdev2 = mdev**2.
COMPUTE cp = fdev*mdev.

FORMAT fverb mverb (f2.0).
LIST fverb mverb fdev fdev2 mdev mdev2 cp.
```



fverb	mverb	fdev	fdev2	mdev	mdev2	cp
15	13	.67	.44	1.67	2.78	1.11
11	10	-3.33	11.11	-1.33	1.78	4.44
17	12	2.67	7.11	.67	.44	1.78
9	10	-5.33	28.44	-1.33	1.78	7.11
13	9	-1.33	1.78	-2.33	5.44	3.11
21	14	6.67	44.44	2.67	7.11	17.78

```
DESCR fdev2 mdev2 cp /STAT = SUM.
```

	N	Sum	
fdev2	6	93.33	***= SSx***
mdev2	6	19.33	***= SSy***
cp	6	35.33	***= SCP***

r = 35.33/SQRT(93.33x19.33) = .8318

```
CORRELATION /VARI = fverb mverb.
```

		FVERB	MVERB
FVERB	Pearson	1	.832
	Sig.	.	.040

t_r = .832/SQRT((1-.832^2)/(6-2)) = 2.999 df = 6 - 2 = 4 $t_{Critical}$=2.132 Rej H_0

F_R = (.8318^2/1)/((1-.8318^2)/(6-2))= 8.982 = t_r^2 df = 1, 4 $F_{Critical}$=4.54

```
*4&5&6&7-Regression.
REGRESS /DEPEND = mverb   /ENTER fverb /SAVE PRED(prdm) RESI(resm).
```

Model	R	R Square	Adjusted R Square	Std. Error Estimate
1	.832(a)	.692	.615	1.2203629

R^2 = 13.376/19.333 = .692 = r
1-R^2 = 5.957/19.333 = .308
√.692 = .832 √.308 = .555

Model		SS	df	Mean Square	F	Sig.
1	Regression	13.376	1	13.376	8.982	.040(a)
	Residual	5.957	4	1.489		
	Total	19.333	5			

MS = SS/df
F = $MS_{Regression}/MS_{Residual}$ = t_r^2

Coefficients(a) Model		Unstandardized B	Std. Error	Standardized Beta	t	Sig.
1	(Constant)	5.907	1.878		3.146	.035
	FVERB	.379	.126	.832	2.997	.040

b_1 = 35.333/93.333
b_0 = 11.333 - .379 x 14.333

SE_{b1} = SQRT(1.489/93.33) t_{b1} = (.379-0)/.126) = t_r

Residuals	Minimum	Maximum	Mean	Std.Dev.	N
Predicted	9.314285	13.857142	11.333333	1.6356155	6
Residual	-1.828571	1.414286	.000000	1.0915258	6

$SS_{Regression} = (6-1)x1.6356155^2 = 13.376$ $SS_{Residual} = (6-1)x1.0915258^2 = 5.957$

LIST fverb mverb prdm resm

	y	*y′*	*y-y′*
fverb	mverb	prdm	resm
15	13	11.58571	1.41429
11	10	10.07143	-.07143
17	12	12.34286	-.34286
9	10	9.31429	.68571
13	9	10.82857	-1.82857
21	14	13.85714	.14286

***Correlate original and derived variables (not done in lab).**
CORR fverb mverb prdm resm /STAT = DESCR.

	Mean	Std. Deviation	N
fverb	14.33	4.320	6
mverb	11.33	1.966	6
prdm	11.3333333	1.63561551	6
resm	.0000000	1.09152580	6

			X	*Y*	*Y′*	*Y-Y′*
Correlations			fverb	mverb	prdm	resm
mverb	Pearson	***Y***	.832			
prdm	Pearson	***Y′***	1.000	.832		
resm	Pearson	***Y-Y′***	.000	.555	.000	

***8-Paired t-test (same data entry as for corr/reg).**

COMPUTE diff = fverb-mverb.
LIST.

fverb	mverb	diff
15.0000	13.0000	2.0000
11.0000	10.0000	1.0000
17.0000	12.0000	5.0000
9.0000	10.0000	-1.0000
13.0000	9.0000	4.0000
21.0000	14.0000	7.0000

$t_D = M_D/(s_D/\sqrt{6})$
$= 3.0/1.1832$
$= 2.535$

TTEST /TESTVALU = 0 /VARI = diff.

	N	Mean	Std. Deviation	Std. Error Mean
diff	6	3.000000	2.8982753	1.1832160

t_{Diff} = 3.0/1.183 = 2.54

$s_D^2 = 4.3205^2 + 1.9664^2 - 2 \times .8318 \times 4.3205 \times 1.9664 = 8.3998$

	t	df	Sig. (2-tailed)	Mean Difference
diff	2.535	5	.052	3.0000000

TTEST PAIR fverb mverb.

		Mean	N	Std. Deviation	Std. Error Mean
Pair	fverb	14.333333	6	4.3204938	1.7638342
	mverb	11.333333	6	1.9663842	.8027730

		N	Correlation	Sig.
Pair 1	fverb & mverb	6	.832	.040

	Paired Differences			t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean			
Pair 1 fverb - mverb	3.0000000	2.8982753	1.1832160	2.535	5	.052

***equivalent to following F-test (not done in class).**

MANOVA diff /PRINT = CELL.

	Mean	Std. Dev.	N
For entire sample	3.000	2.898	6

Source of Variation	SS	DF	MS	F	Sig of F
WITHIN CELLS	42.00	5	8.40		
CONSTANT	54.00	1	54.00	6.43	.052

***9-Independent Groups t-test and F test.**
*** Show correspondences ... very few calculations (most shown below are to practice).**

***Convert from paired scores to independent scores.**

VARSTOCASES /MAKE verbal FROM fverb mverb /INDEX gend.
SORT CASES BY gend.

***Or re-enter data in Independent format.**

```
DATA LIST FREE/ gend verbal.
BEGIN DATA
1 15        1 11        1 17        1  9        1 13        1 21
2 13        2 10        2 12        2 10        2  9        2 14
END DATA.
```

TTEST /GROUP = gend /VARI = verbal.

	gend	N	Mean	Std. Deviation	Std. Error Mean
verbal	1.0000	6	14.333333	4.3204938	1.7638342
	2.0000	6	11.333333	1.9663842	.8027730

S^2_{Pooled} = (93.33+13.33)/(6+6-2) = 11.2667

	~~Levene's Test for Variances~~ F	~~Sig.~~	t-test for Equality of Means t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
verbal Equal variances	~~2.841~~	~~.123~~	1.548	10	.153	3.0000000	1.9379256

S_p^2 = 11.2667 SE = SQRT(11.2667(1/6 + 1/6)) > SE_{PDtest}

GLM verbal BY gend /STAT = DESCR.

	N	Mean	Std. Deviation	Std. Error	$M_j - M_G$
1.0000	6	14.333333	4.3204938	1.7638342	+1.5
2.0000	6	11.333333	1.9663842	.8027730	-1.5
Total	12	12.833333	3.5632807	1.0286305	s_M^2 = (1.5² + -1.5²)/(2-1) = 4.50

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	27.000	1	27.000	2.396	.153
Within Groups	112.667	10	11.267		
Total	139.667	11			

F = 6x4.5/11.267 = 27.0/11.267 = t^2

***Correlation / Regression and Ind groups ttest / anova.**

CORRELATION /VARI = gend verbal.

		GEND	VERBAL
GEND	Pearson	1	-.440
	Sig. (2-tailed)	.	.153

$= Sig_F = Sig_t$ $> .052 = p_{PDttest}$

REGRESS /VARI = gend verbal /DEPEND = verbal /ENTER /SAVE PRED(prdgen) RESI(resgen).

Model	R	R Square	Adjusted R Square	Std. Error of Estimate
1	.440(a)	.193	.113	3.3565856

Model		SS	df	Mean Square	F	Sig.	
1	Regression	**27.000**	**1**	**27.000**	**2.396**	**.153**(a)	$F_{Reg} = F_{Ind}$
	Residual	**112.667**	**10**	**11.267**			
	Total	139.667	11				

Coefficients(a) Model		Unstandardized B	Std. Error	Standardized Beta	t	Sig.
1	(Constant)	17.333	3.064		5.657	.000
	GEND	**-3.000**	**1.938**	-.440	**-1.548**	**.153**
		$M_{Male-Female}$	SE_t		$t_{Gend} = t_{Ind}$	

LIST gend verbal prdgen resgen.

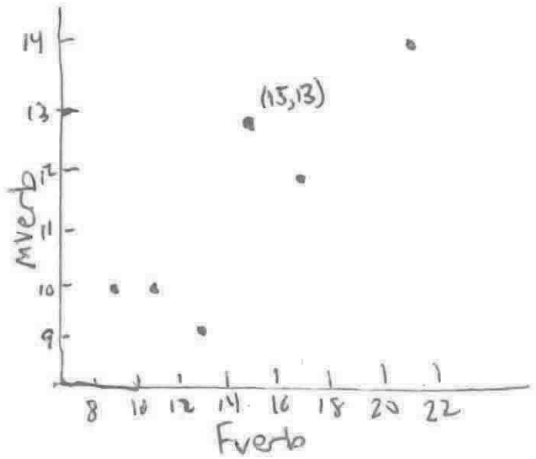
GEND	VERBAL	PRDGEN (*Y′=Ms*)	RESGEN
1.0000	15.0000	14.33333	.66667
1.0000	11.0000	14.33333	-3.33333
1.0000	17.0000	14.33333	2.66667
1.0000	9.0000	14.33333	-5.33333
1.0000	13.0000	14.33333	-1.33333
1.0000	21.0000	14.33333	6.66667
2.0000	13.0000	11.33333	1.66667
2.0000	10.0000	11.33333	-1.33333
2.0000	12.0000	11.33333	.66667
2.0000	10.0000	11.33333	-1.33333
2.0000	9.0000	11.33333	-2.33333
2.0000	14.0000	11.33333	2.66667

DESCR prdgen resgen.

	N	Minimum	Maximum	Mean	Std. Deviation
PRDGEN	12	11.33333	14.33333	12.8333333	1.56669890
RESGEN	12	-5.33333	6.66667	.0000000	3.20037877

1:3 Lab - Correlation + Regression

Q1. Create a rough graph (scatterplot)



Q3. summary of relationship, r Line 7 of formula sheet, need SCP, SS_x SS_y

$\bar{X} = 14.33$ $\bar{Y} = 11.33$

eg. SCP:

X	$(x-\bar{x})$	Y	$(y-\bar{y})$	$(x-\bar{x})(y-\bar{y})$
15	.667	13	1.667	1.111
11	-3.333	10	-1.333	4.444
17	2.667	12	.667	1.778
9	-5.333	10	-1.333	7.111
13	-1.333	9	-2.333	3.111
21	6.667	14	2.667	17.778
$\Sigma x = 86$		$\Sigma y = 68$		SCP = 35.33

$\Sigma(x-\bar{x})(y-\bar{y})$ ← line 6

or $\Sigma xy - \frac{(\Sigma x)(\Sigma y)}{n} = \frac{1010 - (86)(68)}{6} = 1010 - 974.667 = 35.33$

(n ↖ six comparisons; 6 ↗)

$r = \frac{SCP}{\sqrt{SS_x SS_y}} = \frac{35.33}{\sqrt{(93.33)(19.33)}} = \frac{35.33}{42.479} = 0.8318$ ← strong correlation

---

Q4 Predicted Scores for males (based on Female scores) Line 7 formula

$\hat{Y} = b_0 + b_1 X$ → need b_0 and b_1; get b_1 first

$b_1 = \frac{SCP}{SS_x} = \frac{35.33}{93.33} = 0.3786$ slope

$b_0 = \bar{Y} - b_1\bar{X} = 11.33 - (0.3786)(14.33) = 5.907$ y-intercept

Q4 (cont'd)

use line equation to calculate predicted scores

$\hat{Y} = b_0 + b_1 X$

$= 5.907 + 0.3786 X$

X	$b_1 X$	$\hat{Y} = b_0 + b_1 X$
15	5.679	11.586
11	4.164	10.071
17	6.436	12.343
9	3.407	9.314
13	4.921	10.829
21	7.950	13.857

Q6. Two additional tests:

SPSS analysis was F-test - line 12: $F = \frac{r^2/1}{\frac{1-r^2}{n-2}} = \frac{0.6918/1}{\frac{(1-0.6918)}{6-2}} = \frac{.6918}{.3082/4}$

$= \boxed{8.98}$

To confirm, (Line 10)

$t_r = \frac{r - 0}{\sqrt{\frac{1-r^2}{n-2}}}$ ← if null hyp. is true, no correlation ($r = 0$)

← $df = n-2$

$t^2 = F$

$r = 0.8318 \quad r^2 = 0.6919 \quad n = 6$

$t_r = \frac{0.8318 - 0}{\sqrt{\frac{(1-0.6918)}{6-2}}} = \frac{.8318}{\sqrt{\frac{.3082}{4}}} = \frac{.8318}{.2778} = \boxed{2.997}$

Second test (line 11), t_{b_1} for slope

$t_{b_1} = \frac{b_1 - 0}{\sqrt{\frac{MS_{res}}{SS_X}}}$ } standard error; can take from SPSS printout

$t_{b_1} = \frac{0.3786 - 0}{0.126} = 2.997 \simeq 3.00$

Q7 strength of relationship, → r^2

ratio of SS_{reg}/SS_y

Line 8 $r^2 = \frac{SS_{reg}}{SS_y}$

from SPSS printout (Q3+4) - $SS_{reg} = 13.376$

$SS_{res} = 5.957$

$SS_{Tot} = 19.333 = SS_y$

$r^2 = \frac{SS_{reg}}{SS_y} = \frac{13.376}{19.333} = .6918$

$r = \sqrt{r^2} = \sqrt{.6918} = .832$

Q8 paired t-test Line 5

$t = \frac{\bar{D} - 0}{S_D / \sqrt{n_D}}$ $\bar{D} = \bar{X} - \bar{Y} = 14.33 - 11.33 = \boxed{3.00}$ $n_D = \boxed{6}$

$S_D = \sqrt{S^2_D} = \sqrt{S_1^2 + S_2^2 - 2 \times r_{12} \times S_1 \times S_2}$

$= \sqrt{(4.320)^2 + (1.967)^2 - [(2)(.8318)(4.3205)(1.9664)]}$

$= \sqrt{8.399} = \boxed{2.898}$ → S_D on printout

$\frac{3.00 - 0}{2.898/\sqrt{6}} = \frac{3}{1.183} = \boxed{2.535}$

Q9 Independent t-test

$t = \frac{(\bar{Y}_1 - \bar{Y}_2) - 0}{\sqrt{S^2_p \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

$S^2_p = \frac{SS_1 + SS_2}{n_1 + n_2 - 2}$ $SS_1 = 93.33$ $SS_2 = 19.33$

$S^2_p = \frac{93.33 + 19.33}{6 + 6 - 2} = \frac{112.66}{10} = 11.26$

$\bar{Y}_1 = 14.33$ $n_1 = 6$ $S^2_p = ?$

$\bar{Y}_2 = 11.33$ $n_2 = 6$

$t = \frac{14.33 - 11.33 - 0}{\sqrt{(11.266)\left(\frac{1}{6} + \frac{1}{6}\right)}} = \frac{3.00}{\sqrt{3.755}} = \boxed{1.548}$

Q9 (cont'd) Line 4

$$F = \frac{n_j s^2_{\bar{y}}}{s^2_p}$$

n_j = # of obs. per group = 6

$s^2_p = 11.266$

$s^2_{\bar{y}} = ??$

$$s^2_{\bar{y}} = \frac{\Sigma(\bar{y}_j - \bar{y}_G)^2}{K-1}$$

K = # of groups = 2

$\bar{y}_j$ = "group mean" = 14.33, 11.33

$\bar{y}_G$ = "grand mean" = $\frac{(14.33 + 11.33)}{2} = 12.83$

$\bar{y}_j$	$\bar{y}_j - \bar{y}_G$	$(\bar{y}_j - \bar{y}_G)^2$
14.33	1.5	2.25
11.33	-1.5	2.25
		$\Sigma = 4.50$

So, $s^2_{\bar{y}} = \frac{4.50}{2-1} = 4.50$

$$F = \frac{n_j s^2_{\bar{y}}}{s^2_p} = \frac{(6)(4.50)}{11.266} = 2.396 = (1.548)^2$$