

# Parametric Bivariate Tests

with associated Stata code and output

## Pearson Correlation

dependent variable: continuous  
independent variable: continuous

- `pwcorr alc93 age, sig`

	alc93	age
alc93	1.0000	
age	-0.2078	1.0000

## Linear regression (a multivariate technique but used here as bivariate)

dependent variable: continuous  
independent variable: continuous

- `reg alc93 age, beta`

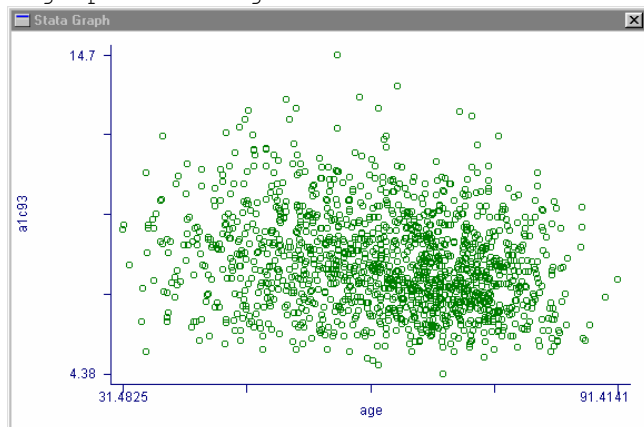
Source	SS	df	MS	Number of obs =	1431
Model	147.473892	1	147.473892	F( 1, 1429) =	64.51
Residual	3266.7938	1429	2.28606984	Prob > F =	0.0000
Total	3414.2677	1430	2.38759979	R-squared =	0.0432
				Adj R-squared =	0.0425
				Root MSE =	1.512

	"beta"				"standardized"
alc93	Coef.	Std. Err.	t	P> t	Beta
age	-.0272079	.0033875	-8.03	0.000	-.2078303
_cons	9.681808	.2184602	44.32	0.000	.

Note that the overall beta coefficient is virtually identical to the correlation coefficient produced by `pwcorr` above.

It's also helpful to graph your two variables (scatter plot) to see what the distribution looks like

`. graph alc93 age`



**T-Test**

dependent variable: continuous  
independent variable: dichotmous

- ttest alc93, by(married)

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	416	8.038205	.0807075	1.646116	7.879558	8.196851
1	995	7.927545	.0473975	1.495088	7.834535	8.020556
combined	1411	7.960171	.0410345	1.54139	7.879675	8.040666
diff		.1106595	.0899785		-.0658468	.2871658

Degrees of freedom: 1409

Ho: mean(0) - mean(1) = diff = 0

Ha: diff < 0	Ha: diff ~= 0	Ha: diff > 0
t = 1.2298	t = 1.2298	t = 1.2298
P < t = 0.8905	P >  t  = 0.2190	P > t = 0.1095

**Analysis of Variance (ANOVA)**

dependent variable: continuous  
independent variable: > 2 categories

- oneway alc93 race, sch

Source	Analysis of Variance			F	Prob > F
	SS	df	MS		
Between groups	12.1166577	2	6.05832887	2.54	0.0790
Within groups	3328.91281	1397	2.38290108		
Total	3341.02947	1399	2.38815545		

Bartlett's test for equal variances: chi2(2) = 7.5251 Prob>chi2 = 0.023

Comparison of alc93 by race  
(Scheffe)

Row Mean-	Col Mean	
	1	2
2	.465299	
	0.084	
3	.08286	-.382439
	0.893	0.353

Use the "oneway" command rather than the "anova" command as it gives the Scheffe option

### Logit function

dependent variable: dichotomous

independent variable: continuous (or categorical)

- logit hisugar2 age

Iteration 0: log likelihood = -1079.1344

Iteration 1: log likelihood = -1064.3064

Iteration 2: log likelihood = -1064.2957

```
Logit estimates                                     Number of obs =      1607
                                                    LR chi2(1)      =      29.68
                                                    Prob > chi2     =      0.0000
Log likelihood = -1064.2957                       Pseudo R2      =      0.0138
```

hisugar2	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
age	-.0235094	.0043502	-5.40	0.000	-.0320357	-.0149832
_cons	1.05084	.2758603	3.81	0.000	.510164	1.591517

*To see odds ratio do a logistic regression (a multivariate technique but used as a bivariate here since only one independent variable is included in the model)*

- logistic hisugar2 age

```
Logit estimates                                     Number of obs =      1607
                                                    LR chi2(1)      =      29.68
                                                    Prob > chi2     =      0.0000
Log likelihood = -1064.2957                       Pseudo R2      =      0.0138
```

hisugar2	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
age	.9767647	.0042491	-5.40	0.000	.968472	.9851285

### Chi-square

dependent variable: categorical

independent variable: categorical

- tab hisugar race, chi

symptoms of high blood sugar	race			Total
	1	2	3	
1	306	11	22	339
2	253	17	16	286
3	146	10	8	164
4	307	17	21	345
5	404	9	27	440
Total	1416	64	94	1574

Pearson chi2(8) = 10.6282 Pr = 0.224

A more detailed output may be obtained by adding the option of seeing row and column percentages.

- tab hisugar race, ro col chi

symptoms	race			Total
of high	1	2	3	
blood				
sugar	1	2	3	Total
1	306	11	22	339
	90.27	3.24	6.49	100.00
	21.61	17.19	23.40	21.54
2	253	17	16	286
	88.46	5.94	5.59	100.00
	17.87	26.56	17.02	18.17
3	146	10	8	164
	89.02	6.10	4.88	100.00
	10.31	15.63	8.51	10.42
4	307	17	21	345
	88.99	4.93	6.09	100.00
	21.68	26.56	22.34	21.92
5	404	9	27	440
	91.82	2.05	6.14	100.00
	28.53	14.06	28.72	27.95
Total	1416	64	94	1574
	89.96	4.07	5.97	100.00
	100.00	100.00	100.00	100.00

Pearson chi2(8) = 10.6282 Pr = 0.224