

Sta442/1008f05 Overheads 3: Explore math data with elementary tests

```
/* mathtests.sas */
%include 'mathexread2.sas';
title2 'Explore math data with elementary tests';

options pagesize=1000; /* Print fewer page headings */

proc freq;
  title3 'Chisquared test of independence';
  tables (sex tongue ethnic course) * passed / chisq nocol nopercnt;
  /* These options give just observed frequency and percentage
     of row total in each cell. We want to look at percent who
     passed the course */

/* In the course*passed chisquare, no student passed the course if she or he
failed to respond to the question about what course they were in (they didn't
know, probably). We would like to see if there is still a significant
relationship if you set these folks aside. So, the following shows how to do
analyses on a subset of the data. */

data mathex2;
  set mathex;
  if course ne 4; /* Keep just these cases */

proc freq;
  title3 'Course*passed without course=unknown';
  tables course * passed / chisq nocol nopercnt;

/* SAS procedures use the most recently created data set by default. That's
why we need the data=mathex option in the next procedure. */

proc ttest data=mathex;
  title3 'Independent t-test';
  class sex;
  var gpa english hscalcalc totscore grade;

proc ttest data=mathex;
  title3 'Independent t-test';
  class tongue;
  var gpa english hscalcalc totscore grade;

/* Now illustrate how to do a matched t-test. In the diagnostic test, compare
performance on the precalculus material to performance on the calculus
material. There are 9 precalc and 11 calc. Test whether difference in
PERCENTAGE right is different from zero. */
```

```

data mathex3;
  set mathex;
  pcalcper = 100*precalc/9;
  calcper = 100*calc/11;
  diff = pcalcper - calcper;
  /* Positive numbers mean calculus questions are harder */

proc means n mean std t probt data = mathex3;
  title3 'Matched t-test: Precalc vs. calc on diagnostic test';
  var pcalcper calcper diff;

proc glm data = mathex;
  title3 'One-way anova';
  class ethnic;
  model gpa english hscalc totscore grade = ethnic;
  means ethnic;
  means ethnic / Bon;

options pagesize = 35;

proc plot data = mathex;
  title3 'Scatterplot';
  plot grade * totscore;

proc reg data = mathex;
  title3 'Simple regression';
  model grade = totscore;

proc corr data = mathex;
  title3 'Correlation Matrix';
  var precalc calc grade;

/* Labels (usually handy) are too long for a big correlation matrix. Here is a
work-around */

data mathex4;
  set mathex;
  label
    tongue = ' '
    gpa = ' '
    english = ' '
    finmat = ' '
    alggeo = ' '
    hscalc = ' '
    nhsmath = ' '
    precalc = ' '
    calc = ' '
    totscore = ' '
    grade = ' ';

options pagesize=200;

proc corr;
  title3 'Bigger Correlation Matrix';
  var gpa english finmat alggeo hscalc nhsmath precalc calc
    totscore grade tongue;

```

tuzo > cat mathtests.lst

Math Diagnostic Study: Exploratory data 1
Explore math data with elementary tests
Chisquared test of independence
21:49 Thursday, September 1, 2005

The FREQ Procedure

Table of sex by passed

sex	passed(Passed the course)		Total
	No	Yes	
Female	118 44.36	148 55.64	266
Male	138 48.42	147 51.58	285
Total	256	295	551

Frequency Missing = 28

Statistics for Table of sex by passed

Statistic	DF	Value	Prob
Chi-Square	1	0.9118	0.3396
Likelihood Ratio Chi-Square	1	0.9122	0.3395
Continuity Adj. Chi-Square	1	0.7559	0.3846
Mantel-Haenszel Chi-Square	1	0.9101	0.3401
Phi Coefficient		-0.0407	
Contingency Coefficient		0.0406	
Cramer's V		-0.0407	

Fisher's Exact Test

Cell (1,1) Frequency (F)	118
Left-sided Pr <= F	0.1923
Right-sided Pr >= F	0.8509
Table Probability (P)	0.0432
Two-sided Pr <= P	0.3484

Effective Sample Size = 551
Frequency Missing = 28

Table of tongue by passed

tongue (Mother Tongue)	passed (Passed the course)		Total
	No	Yes	
English	187 46.52	215 53.48	402
Other	69 46.31	80 53.69	149
Total	256	295	551

Frequency Missing = 28

Statistics for Table of tongue by passed

Statistic	DF	Value	Prob
Chi-Square	1	0.0019	0.9652
Likelihood Ratio Chi-Square	1	0.0019	0.9652
Continuity Adj. Chi-Square	1	0.0000	1.0000
Mantel-Haenszel Chi-Square	1	0.0019	0.9652
Phi Coefficient		0.0019	
Contingency Coefficient		0.0019	
Cramer's V		0.0019	

Fisher's Exact Test

Cell (1,1) Frequency (F)	187
Left-sided Pr <= F	0.5552
Right-sided Pr >= F	0.5214
Table Probability (P)	0.0765
Two-sided Pr <= P	1.0000

Effective Sample Size = 551

Frequency Missing = 28

Table of ethnic by passed

ethnic(Apparent ethnic background based on name (compromise))			
passed(Passed the course)			
Frequency	No	Yes	Total
Row Pct			
Asian	65 49.62	66 50.38	131
European	118 45.74	140 54.26	258
Middle-Eastern	32 45.07	39 54.93	71
East Indian	31 39.74	47 60.26	78
Other or unknown	28 68.29	13 31.71	41
Total	274	305	579

Statistics for Table of ethnic by passed

Statistic	DF	Value	Prob
Chi-Square	4	9.7117	0.0456
Likelihood Ratio Chi-Square	4	9.8394	0.0432
Mantel-Haenszel Chi-Square	1	0.4387	0.5077
Phi Coefficient		0.1295	
Contingency Coefficient		0.1284	
Cramer's V		0.1295	

Sample Size = 579

Table of course by passed

course(Which university calculus course)
 passed(Passed the course)

Frequency	No	Yes	Total
1	48	17	65
Row Pct	73.85	26.15	
2	171	258	429
Row Pct	39.86	60.14	
3	6	30	36
Row Pct	16.67	83.33	
No Resp	49	0	49
Row Pct	100.00	0.00	
Total	274	305	579

Statistics for Table of course by passed

Statistic	DF	Value	Prob
Chi-Square	3	96.0435	<.0001
Likelihood Ratio Chi-Square	3	116.9037	<.0001
Mantel-Haenszel Chi-Square	1	7.5589	0.0060
Phi Coefficient		0.4073	
Contingency Coefficient		0.3772	
Cramer's V		0.4073	

Sample Size = 579

The FREQ Procedure

Table of course by passed

course(Which university calculus course)		passed(Passed the course)		Total
Frequency				
Row	Pct	No	Yes	
1		48	17	65
		73.85	26.15	
2		171	258	429
		39.86	60.14	
3		6	30	36
		16.67	83.33	
Total		225	305	530

Statistics for Table of course by passed

Statistic	DF	Value	Prob
Chi-Square	2	37.2001	<.0001
Likelihood Ratio Chi-Square	2	38.5141	<.0001
Mantel-Haenszel Chi-Square	1	36.2232	<.0001
Phi Coefficient		0.2649	
Contingency Coefficient		0.2561	
Cramer's V		0.2649	

Sample Size = 530

Math Diagnostic Study: Exploratory data
 Explore math data with elementary tests
 Independent t-test

3

21:49 Thursday, September 1, 2005

The TTEST Procedure

Statistics

Variable	sex	N	Lower CL Mean	Mean	Upper CL Mean	Lower CL Std Dev	Std Dev
gpa	Female	220	78.91	79.68	80.449	5.2959	5.7912
gpa	Male	238	78.184	78.976	79.769	5.6955	6.2076
gpa	Diff (1-2)		-0.402	0.7031	1.8079	5.6451	6.0112
english	Female	225	76.941	77.893	78.846	6.635	7.2486
english	Male	245	72.95	74.143	75.336	8.711	9.4829
english	Diff (1-2)		2.2105	3.7505	5.2904	7.9766	8.4872
hscal	Female	215	73.318	74.963	76.608	11.178	12.235
hscal	Male	225	74.363	75.964	77.566	11.156	12.188
hscal	Diff (1-2)		-3.29	-1.002	1.2872	11.453	12.211
totscore	Female	237	6.7169	7.1392	7.5616	3.0275	3.3002
totscore	Male	231	7.8269	8.3377	8.8484	3.6101	3.9396
totscore	Diff (1-2)		-1.858	-1.198	-0.539	3.4111	3.6299
grade	Female	193	55.513	58.13	60.746	16.758	18.432
grade	Male	190	56.306	59.205	62.105	18.407	20.26
grade	Diff (1-2)		-4.966	-1.076	2.8146	18.078	19.36

Statistics

Variable	sex	Upper CL Std Dev	Std Err	Minimum	Maximum
gpa	Female	6.3895	0.3904	66	94
gpa	Male	6.8216	0.4024	65	97.3
gpa	Diff (1-2)	6.4285	0.5622		
english	Female	7.9882	0.4832	53	92
english	Male	10.406	0.6058	50	96
english	Diff (1-2)	9.0682	0.7837		
hscal	Female	13.515	0.8344	50	99
hscal	Male	13.431	0.8125	50	100
hscal	Diff (1-2)	13.077	1.1646		
totscore	Female	3.6274	0.2144	1	17
totscore	Male	4.3357	0.2592	0	20
totscore	Diff (1-2)	3.8789	0.3356		
grade	Female	20.48	1.3268	4	97
grade	Male	22.531	1.4698	1	99
grade	Diff (1-2)	20.84	1.9786		

T-Tests					
Variable	Method	Variances	DF	t Value	Pr > t
gpa	Pooled	Equal	456	1.25	0.2117
gpa	Satterthwaite	Unequal	456	1.25	0.2105
english	Pooled	Equal	468	4.79	<.0001
english	Satterthwaite	Unequal	453	4.84	<.0001
hscal	Pooled	Equal	438	-0.86	0.3902
hscal	Satterthwaite	Unequal	437	-0.86	0.3902
totscore	Pooled	Equal	466	-3.57	0.0004
totscore	Satterthwaite	Unequal	448	-3.56	0.0004
grade	Pooled	Equal	381	-0.54	0.5870
grade	Satterthwaite	Unequal	376	-0.54	0.5873

Equality of Variances					
Variable	Method	Num DF	Den DF	F Value	Pr > F
gpa	Folded F	237	219	1.15	0.2969
english	Folded F	244	224	1.71	<.0001
hscal	Folded F	214	224	1.01	0.9535
totscore	Folded F	230	236	1.42	0.0070
grade	Folded F	189	192	1.21	0.1926

Now independent variable = tongue

Math Diagnostic Study: Exploratory data
 Explore math data with elementary tests
 Independent t-test

4

21:49 Thursday, September 1, 2005

The TTEST Procedure

Statistics

Variable	tongue	N	Lower CL Mean	Mean	Upper CL Mean	Lower CL Std Dev	Std Dev
gpa	English	348	78.787	79.422	80.056	5.6024	6.0188
gpa	Other	110	77.837	78.974	80.111	5.3136	6.0174
gpa	Diff (1-2)		-0.846	0.4482	1.7419	5.6519	6.0185
english	English	354	76.161	76.994	77.828	7.4271	7.9745
english	Other	116	70.891	72.716	74.54	8.7857	9.9186
english	Diff (1-2)		2.4933	4.2788	6.0644	7.9825	8.4935
hscal	English	330	73.583	74.894	76.205	11.25	12.109
hscal	Other	110	74.877	77.218	79.56	10.942	12.391
hscal	Diff (1-2)		-4.96	-2.324	0.3112	11.424	12.18
totscore	English	346	7.1858	7.5751	7.9644	3.4263	3.6817
totscore	Other	122	7.5205	8.1721	8.8238	3.2296	3.6357
totscore	Diff (1-2)		-1.356	-0.597	0.1623	3.4486	3.6698
grade	English	289	55.112	57.339	59.566	17.783	19.234
grade	Other	94	58.8	62.734	66.668	16.8	19.208
grade	Diff (1-2)		-9.884	-5.395	-0.906	17.954	19.227

Statistics

Variable	tongue	Upper CL Std Dev	Std Err	Minimum	Maximum
gpa	English	6.5026	0.3226	66	96.2
gpa	Other	6.9377	0.5737	65	97.3
gpa	Diff (1-2)	6.4362	0.6583		
english	English	8.6096	0.4238	50	96
english	Other	11.39	0.9209	50	93
english	Diff (1-2)	9.075	0.9087		
hscal	English	13.111	0.6666	50	100
hscal	Other	14.286	1.1814	50	98
hscal	Diff (1-2)	13.043	1.3409		
totscore	English	3.9786	0.1979	0	20
totscore	Other	4.1594	0.3292	1	19
totscore	Diff (1-2)	3.9216	0.3864		
grade	English	20.944	1.1314	2	99
grade	Other	22.428	1.9811	1	95
grade	Diff (1-2)	20.697	2.283		

T-Tests					
Variable	Method	Variances	DF	t Value	Pr > t
gpa	Pooled	Equal	456	0.68	0.4963
gpa	Satterthwaite	Unequal	183	0.68	0.4968
english	Pooled	Equal	468	4.71	<.0001
english	Satterthwaite	Unequal	166	4.22	<.0001
hscalcalc	Pooled	Equal	438	-1.73	0.0837
hscalcalc	Satterthwaite	Unequal	183	-1.71	0.0883
totscore	Pooled	Equal	466	-1.54	0.1230
totscore	Satterthwaite	Unequal	214	-1.55	0.1216
grade	Pooled	Equal	381	-2.36	0.0186
grade	Satterthwaite	Unequal	158	-2.36	0.0193

Equality of Variances					
Variable	Method	Num DF	Den DF	F Value	Pr > F
gpa	Folded F	347	109	1.00	1.0000
english	Folded F	115	353	1.55	0.0027
hscalcalc	Folded F	109	329	1.05	0.7479
totscore	Folded F	345	121	1.03	0.8855
grade	Folded F	288	93	1.00	1.0000

Math Diagnostic Study: Exploratory data 5
 Explore math data with elementary tests
 Matched t-test: Precalc vs. calc on diagnostic test
 21:49 Thursday, September 1, 2005

The MEANS Procedure					
Variable	N	Mean	Std Dev	t Value	Pr > t
pcalcper	480	48.9120370	19.5795776	54.73	<.0001
calcper	480	30.1704545	22.6530626	29.18	<.0001
diff	480	18.7415825	21.2553367	19.32	<.0001

Math Diagnostic Study: Exploratory data 6
 Explore math data with elementary tests
 One-way anova
 21:49 Thursday, September 1, 2005

The GLM Procedure

Class Level Information

Class	Levels	Values
ethnic	5	Asian East Indian European Middle-Eastern Other or unknown

Number of observations 579

Dependent Variables With
 Equivalent Missing Value Patterns

Pattern	Obs	Dependent Variables
1	466	gpa
2	480	english
3	448	hscal
4	480	totscore
5	393	grade

NOTE: Variables in each group are consistent with respect to the presence or absence of missing values.

Math Diagnostic Study: Exploratory data 7
 Explore math data with elementary tests
 One-way anova
 21:49 Thursday, September 1, 2005

The GLM Procedure

Dependent Variable: gpa High School GPA

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	267.81815	66.95454	1.85	0.1182
Error	461	16687.65687	36.19882		
Corrected Total	465	16955.47502			

R-Square	Coeff Var	Root MSE	gpa Mean
0.015795	7.590027	6.016546	79.26910

Source	DF	Type I SS	Mean Square	F Value	Pr > F
ethnic	4	267.8181542	66.9545385	1.85	0.1182

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ethnic	4	267.8181542	66.9545385	1.85	0.1182

Math Diagnostic Study: Exploratory data 8
 Explore math data with elementary tests
 One-way anova
 21:49 Thursday, September 1, 2005

The GLM Procedure

Level of ethnic	N	Mean	Std Dev
Asian	107	78.6420561	6.21146289
East Indian	70	80.7300000	5.90333607
European	222	79.3527027	5.74896150
Middle-Eastern	53	78.8207547	6.82124392
Other or unknown	14	77.1285714	6.01043049

No need to test which means are significantly different from which other means. Skip the follow-up tests.

Math Diagnostic Study: Exploratory data 11
 Explore math data with elementary tests
 One-way anova

21:49 Thursday, September 1, 2005

The GLM Procedure

Dependent Variable: english Mark in HS English

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	2706.60022	676.65006	9.54	<.0001
Error	475	33699.50344	70.94632		
Corrected Total	479	36406.10367			

R-Square	Coeff Var	Root MSE	english Mean
0.074345	11.10684	8.422964	75.83583

Source	DF	Type I SS	Mean Square	F Value	Pr > F
ethnic	4	2706.600223	676.650056	9.54	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ethnic	4	2706.600223	676.650056	9.54	<.0001

Math Diagnostic Study: Exploratory data 12
 Explore math data with elementary tests
 One-way anova

21:49 Thursday, September 1, 2005

The GLM Procedure

Level of ethnic	N	-----english----- Mean	Std Dev
Asian	110	71.8363636	10.0416813
East Indian	73	78.1095890	7.9101932
European	224	77.3258929	7.7364694
Middle-Eastern	57	75.3157895	8.0913488
Other or unknown	16	73.9500000	8.9061776

This time we need followups.

The GLM Procedure

Bonferroni (Dunn) t Tests for english

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than Tukey's for all pairwise comparisons.

Alpha 0.05
 Error Degrees of Freedom 475
 Error Mean Square 70.94632
 Critical Value of t 2.82021

Comparisons significant at the 0.05 level are indicated by ***.

ethnic Comparison	Difference Between Means	Simultaneous 95% Confidence Limits	
East Indian - European	0.7837	-2.4177 3.9851	
East Indian - Middle-Eastern	2.7938	-1.4049 6.9925	
East Indian - Other or unknown	4.1596	-2.3976 10.7168	
East Indian - Asian	6.2732	2.6872 9.8593	***
European - East Indian	-0.7837	-3.9851 2.4177	
European - Middle-Eastern	2.0101	-1.5139 5.5341	
European - Other or unknown	3.3759	-2.7712 9.5230	
European - Asian	5.4895	2.7239 8.2552	***
Middle-Eastern - East Indian	-2.7938	-6.9925 1.4049	
Middle-Eastern - European	-2.0101	-5.5341 1.5139	
Middle-Eastern - Other or unknown	1.3658	-5.3548 8.0864	
Middle-Eastern - Asian	3.4794	-0.3974 7.3562	
Other or unknown - East Indian	-4.1596	-10.7168 2.3976	
Other or unknown - European	-3.3759	-9.5230 2.7712	
Other or unknown - Middle-Eastern	-1.3658	-8.0864 5.3548	
Other or unknown - Asian	2.1136	-4.2422 8.4695	
Asian - East Indian	-6.2732	-9.8593 -2.6872	***
Asian - European	-5.4895	-8.2552 -2.7239	***
Asian - Middle-Eastern	-3.4794	-7.3562 0.3974	
Asian - Other or unknown	-2.1136	-8.4695 4.2422	

Math Diagnostic Study: Exploratory data 15
 Explore math data with elementary tests
 One-way anova

21:49 Thursday, September 1, 2005

The GLM Procedure

Dependent Variable: hscalc Mark in HS Calculus

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	1525.80684	381.45171	2.60	0.0354
Error	443	64916.79807	146.53905		
Corrected Total	447	66442.60491			

R-Square	Coeff Var	Root MSE	hscalc Mean
0.022964	16.04541	12.10533	75.44420

Source	DF	Type I SS	Mean Square	F Value	Pr > F
ethnic	4	1525.806845	381.451711	2.60	0.0354

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ethnic	4	1525.806845	381.451711	2.60	0.0354

Math Diagnostic Study: Exploratory data 16
 Explore math data with elementary tests
 One-way anova

21:49 Thursday, September 1, 2005

The GLM Procedure

Level of ethnic	N	Mean	Std Dev
Asian	109	77.3302752	11.8673650
East Indian	68	76.4852941	13.1347695
European	205	73.8439024	12.0189811
Middle-Eastern	52	77.6538462	11.8452185
Other or unknown	14	70.9285714	10.8235092

The GLM Procedure

Bonferroni (Dunn) t Tests for hscalc

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than Tukey's for all pairwise comparisons.

Alpha 0.05
 Error Degrees of Freedom 443
 Error Mean Square 146.539
 Critical Value of t 2.82117

Comparisons significant at the 0.05 level are indicated by ***.

ethnic Comparison	Difference Between Means	Simultaneous 95% Confidence Limits
Middle-Eastern - Asian	0.324	-5.432 6.079
Middle-Eastern - East Indian	1.169	-5.123 7.460
Middle-Eastern - European	3.810	-1.493 9.113
Middle-Eastern - Other or unknown	6.725	-3.558 17.008
Asian - Middle-Eastern	-0.324	-6.079 5.432
Asian - East Indian	0.845	-4.432 6.122
Asian - European	3.486	-0.562 7.535
Asian - Other or unknown	6.402	-3.294 16.097
East Indian - Middle-Eastern	-1.169	-7.460 5.123
East Indian - Asian	-0.845	-6.122 4.432
East Indian - European	2.641	-2.138 7.421
East Indian - Other or unknown	5.557	-4.466 15.580
European - Middle-Eastern	-3.810	-9.113 1.493
European - Asian	-3.486	-7.535 0.562
European - East Indian	-2.641	-7.421 2.138
European - Other or unknown	2.915	-6.518 12.349
Other or unknown - Middle-Eastern	-6.725	-17.008 3.558
Other or unknown - Asian	-6.402	-16.097 3.294
Other or unknown - East Indian	-5.557	-15.580 4.466
Other or unknown - European	-2.915	-12.349 6.518

Ouch.

Math Diagnostic Study: Exploratory data 19
 Explore math data with elementary tests
 One-way anova

21:49 Thursday, September 1, 2005

The GLM Procedure

Dependent Variable: totscore Total # right on diagnostic test

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	219.288140	54.822035	4.10	0.0028
Error	475	6351.303527	13.371165		
Corrected Total	479	6570.591667			

R-Square Coeff Var Root MSE totscore Mean
 0.033374 47.36096 3.656660 7.720833

Source	DF	Type I SS	Mean Square	F Value	Pr > F
ethnic	4	219.2881399	54.8220350	4.10	0.0028

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ethnic	4	219.2881399	54.8220350	4.10	0.0028

Math Diagnostic Study: Exploratory data 20
 Explore math data with elementary tests
 One-way anova

21:49 Thursday, September 1, 2005

The GLM Procedure

Level of ethnic	N	Mean	Std Dev
Asian	109	8.61467890	3.80523094
East Indian	66	8.36363636	3.81751576
European	224	7.08928571	3.43564426
Middle-Eastern	60	8.00000000	3.83118338
Other or unknown	21	7.00000000	4.13521463

The GLM Procedure

Bonferroni (Dunn) t Tests for totscore

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than Tukey's for all pairwise comparisons.

Alpha 0.05
 Error Degrees of Freedom 475
 Error Mean Square 13.37117
 Critical Value of t 2.82021

Comparisons significant at the 0.05 level are indicated by ***.

ethnic Comparison	Difference Between Means	Simultaneous 95% Confidence Limits	
Asian - East Indian	0.2510	-1.3574 1.8595	
Asian - Middle-Eastern	0.6147	-1.0431 2.2724	
Asian - European	1.5254	0.3210 2.7297	***
Asian - Other or unknown	1.6147	-0.8429 4.0723	
East Indian - Asian	-0.2510	-1.8595 1.3574	
East Indian - Middle-Eastern	0.3636	-1.4759 2.2032	
East Indian - European	1.2744	-0.1700 2.7187	
East Indian - Other or unknown	1.3636	-1.2201 3.9473	
Middle-Eastern - Asian	-0.6147	-2.2724 1.0431	
Middle-Eastern - East Indian	-0.3636	-2.2032 1.4759	
Middle-Eastern - European	0.9107	-0.5884 2.4098	
Middle-Eastern - Other or unknown	1.0000	-1.6147 3.6147	
European - Asian	-1.5254	-2.7297 -0.3210	***
European - East Indian	-1.2744	-2.7187 0.1700	
European - Middle-Eastern	-0.9107	-2.4098 0.5884	
European - Other or unknown	0.0893	-2.2642 2.4428	
Other or unknown - Asian	-1.6147	-4.0723 0.8429	
Other or unknown - East Indian	-1.3636	-3.9473 1.2201	
Other or unknown - Middle-Eastern	-1.0000	-3.6147 1.6147	
Other or unknown - European	-0.0893	-2.4428 2.2642	

Math Diagnostic Study: Exploratory data 23
 Explore math data with elementary tests
 One-way anova

21:49 Thursday, September 1, 2005

The GLM Procedure

Dependent Variable: grade Mark in university calculus

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	3631.5326	907.8832	2.50	0.0418
Error	388	140622.4724	362.4291		
Corrected Total	392	144254.0051			

R-Square	Coeff Var	Root MSE	grade Mean
0.025175	32.41104	19.03757	58.73791

Source	DF	Type I SS	Mean Square	F Value	Pr > F
ethnic	4	3631.532642	907.883161	2.50	0.0418

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ethnic	4	3631.532642	907.883161	2.50	0.0418

Math Diagnostic Study: Exploratory data 24
 Explore math data with elementary tests
 One-way anova

21:49 Thursday, September 1, 2005

The GLM Procedure

Level of ethnic	N	Mean	Std Dev
Asian	87	60.0574713	20.9314253
East Indian	53	65.1886792	18.5317364
European	188	56.1542553	18.4243687
Middle-Eastern	50	59.3600000	19.9691190
Other or unknown	15	58.6000000	12.1526011

The GLM Procedure

Bonferroni (Dunn) t Tests for grade

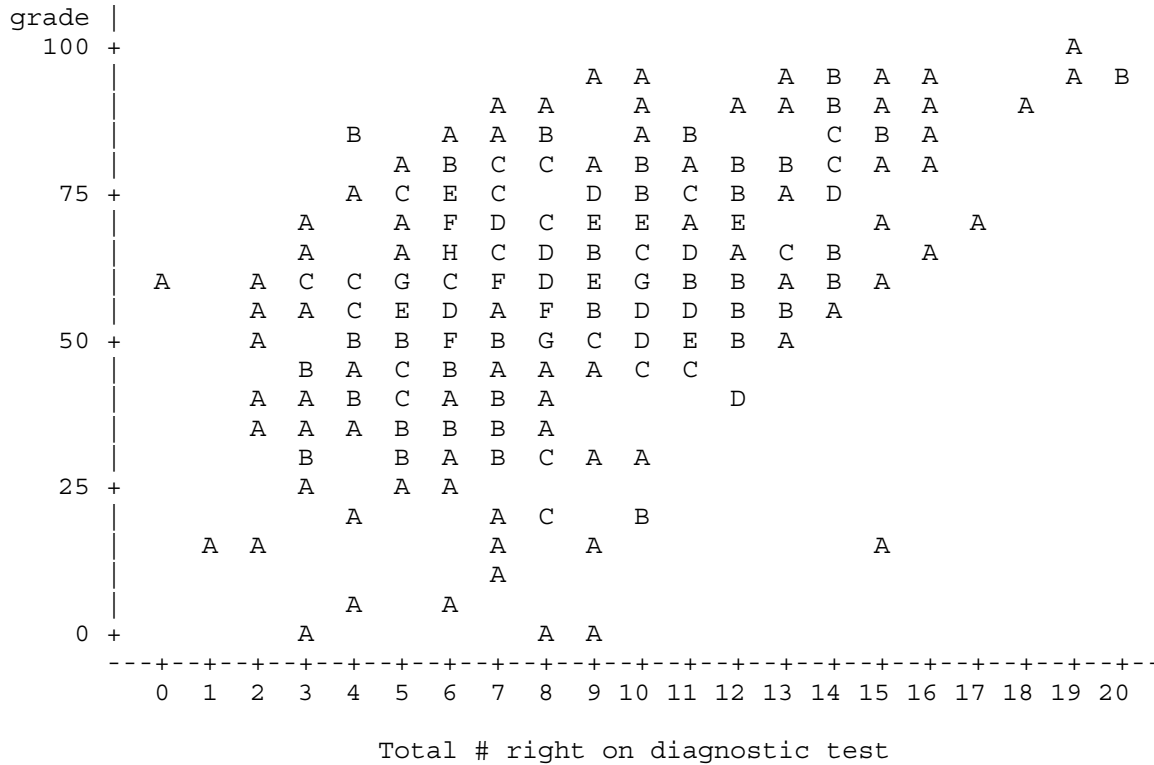
NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than Tukey's for all pairwise comparisons.

Alpha 0.05
 Error Degrees of Freedom 388
 Error Mean Square 362.4291
 Critical Value of t 2.82318

Comparisons significant at the 0.05 level are indicated by ***.

ethnic Comparison	Difference Between Means	Simultaneous 95% Confidence Limits	
East Indian - Asian	5.131	-4.234 14.496	
East Indian - Middle-Eastern	5.829	-4.767 16.425	
East Indian - Other or unknown	6.589	-9.130 22.308	
East Indian - European	9.034	0.676 17.393	***
Asian - East Indian	-5.131	-14.496 4.234	
Asian - Middle-Eastern	0.697	-8.841 10.236	
Asian - Other or unknown	1.457	-13.569 16.484	
Asian - European	3.903	-3.066 10.872	
Middle-Eastern - East Indian	-5.829	-16.425 4.767	
Middle-Eastern - Asian	-0.697	-10.236 8.841	
Middle-Eastern - Other or unknown	0.760	-15.063 16.583	
Middle-Eastern - European	3.206	-5.346 11.758	
Other or unknown - East Indian	-6.589	-22.308 9.130	
Other or unknown - Asian	-1.457	-16.484 13.569	
Other or unknown - Middle-Eastern	-0.760	-16.583 15.063	
Other or unknown - European	2.446	-11.975 16.866	
European - East Indian	-9.034	-17.393 -0.676	***
European - Asian	-3.903	-10.872 3.066	
European - Middle-Eastern	-3.206	-11.758 5.346	
European - Other or unknown	-2.446	-16.866 11.975	

Plot of grade*totscore. Legend: A = 1 obs, B = 2 obs, etc.



NOTE: 233 obs had missing values.

Math Diagnostic Study: Exploratory data 28
 Explore math data with elementary tests
 Simple regression
 21:49 Thursday, September 1, 2005

The REG Procedure
 Model: MODEL1
 Dependent Variable: grade Mark in university calculus

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	23019	23019	77.35	<.0001
Error	344	102369	297.58385		
Corrected Total	345	125388			

Root MSE	17.25062	R-Square	0.1836
Dependent Mean	59.54335	Adj R-Sq	0.1812
Coeff Var	28.97153		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error
Intercept	Intercept	1	40.13541	2.39365
totscore	Total # right on diagnostic test	1	2.26940	0.25803

Math Diagnostic Study: Exploratory data 29
 Explore math data with elementary tests
 Simple regression
 21:49 Thursday, September 1, 2005

The REG Procedure
 Model: MODEL1
 Dependent Variable: grade Mark in university calculus

Parameter Estimates

Variable	Label	DF	t Value	Pr > t
Intercept	Intercept	1	16.77	<.0001
totscore	Total # right on diagnostic test	1	8.80	<.0001

Math Diagnostic Study: Exploratory data
 Explore math data with elementary tests
 Correlation Matrix

30

21:49 Thursday, September 1, 2005

The CORR Procedure

3 Variables: precalc calc grade

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
precalc	480	4.40208	1.76216	2113	0	9.00000
calc	480	3.31875	2.49184	1593	0	11.00000
grade	393	58.73791	19.18319	23084	1.00000	99.00000

Simple Statistics

Variable	Label
precalc	Number precalculus correct
calc	Number calculus correct
grade	Mark in university calculus

Math Diagnostic Study: Exploratory data
 Explore math data with elementary tests
 Correlation Matrix

31

21:49 Thursday, September 1, 2005

The CORR Procedure

Pearson Correlation Coefficients

Prob > |r| under H0: Rho=0

Number of Observations

	precalc	calc	grade
precalc Number precalculus correct	1.00000	0.50135	0.37834
		<.0001	<.0001
	480	480	346
calc Number calculus correct	0.50135	1.00000	0.36247
	<.0001		<.0001
	480	480	346
grade Mark in university calculus	0.37834	0.36247	1.00000
	<.0001	<.0001	
	346	346	393

Math Diagnostic Study: Exploratory data
 Explore math data with elementary tests
 Bigger Correlation Matrix

The CORR Procedure

11 Variables:	gpa	english	finmat	alggeo	hscal	nhsmath
	precalc	calc	totscore	grade	tongue	
Pearson Correlation Coefficients						
Prob > r under H0: Rho=0						
Number of Observations						
	gpa	english	finmat	alggeo	hscal	nhsmath
gpa	1.00000 <.0001 466	0.54327 <.0001 464	0.59057 <.0001 267	0.60194 <.0001 337	0.62249 <.0001 437	0.25305 <.0001 466
english	0.54327 <.0001 464	1.00000 480	0.12663 0.0358 275	0.12841 0.0180 339	0.08498 0.0749 440	0.03698 0.4189 480
finmat	0.59057 <.0001 267	0.12663 0.0358 275	1.00000 284	0.65279 <.0001 176	0.49504 <.0001 254	0.36247 <.0001 284
alggeo	0.60194 <.0001 337	0.12841 0.0180 339	0.65279 <.0001 176	1.00000 347	0.61303 <.0001 341	0.20450 0.0001 347
hscal	0.62249 <.0001 437	0.08498 0.0749 440	0.49504 <.0001 254	0.61303 <.0001 341	1.00000 448	0.30875 <.0001 448
nhsmath	0.25305 <.0001 466	0.03698 0.4189 480	0.36247 <.0001 284	0.20450 0.0001 347	0.30875 <.0001 448	1.00000 579
precalc	0.33965 <.0001 396	0.06543 0.1893 404	0.37257 <.0001 233	0.34221 <.0001 303	0.36684 <.0001 384	0.27452 <.0001 480
calc	0.33876 <.0001 396	0.04272 0.3917 404	0.30320 <.0001 233	0.30966 <.0001 303	0.43774 <.0001 384	0.27881 <.0001 480
totscore	0.39002 <.0001 396	0.05995 0.2293 404	0.38164 <.0001 233	0.37498 <.0001 303	0.47262 <.0001 384	0.31820 <.0001 480
grade	0.58129 <.0001 337	0.16441 0.0022 345	0.65320 <.0001 197	0.60423 <.0001 276	0.53272 <.0001 332	0.14050 0.0053 393
tongue	-0.03187 0.4963 458	-0.21269 <.0001 470	0.02851 0.6360 278	0.07835 0.1476 343	0.08254 0.0837 440	-0.07421 0.0818 551

Pearson Correlation Coefficients
 Prob > |r| under H0: Rho=0
 Number of Observations

	precalc	calc	totscore	grade	tongue
gpa	0.33965 <.0001 396	0.33876 <.0001 396	0.39002 <.0001 396	0.58129 <.0001 337	-0.03187 0.4963 458
english	0.06543 0.1893 404	0.04272 0.3917 404	0.05995 0.2293 404	0.16441 0.0022 345	-0.21269 <.0001 470
finmat	0.37257 <.0001 233	0.30320 <.0001 233	0.38164 <.0001 233	0.65320 <.0001 197	0.02851 0.6360 278
alggeo	0.34221 <.0001 303	0.30966 <.0001 303	0.37498 <.0001 303	0.60423 <.0001 276	0.07835 0.1476 343
hscal	0.36684 <.0001 384	0.43774 <.0001 384	0.47262 <.0001 384	0.53272 <.0001 332	0.08254 0.0837 440
nhsmath	0.27452 <.0001 480	0.27881 <.0001 480	0.31820 <.0001 480	0.14050 0.0053 393	-0.07421 0.0818 551
precalc	1.00000 480	0.50135 <.0001 480	0.81309 <.0001 480	0.37834 <.0001 346	0.07557 0.1025 468
calc	0.50135 <.0001 480	1.00000 480	0.91133 <.0001 480	0.36247 <.0001 346	0.05266 0.2556 468
totscore	0.81309 <.0001 480	0.91133 <.0001 480	1.00000 480	0.42847 <.0001 346	0.07139 0.1230 468
grade	0.37834 <.0001 346	0.36247 <.0001 346	0.42847 <.0001 346	1.00000 393	0.12019 0.0186 383
tongue	0.07557 0.1025 468	0.05266 0.2556 468	0.07139 0.1230 468	0.12019 0.0186 383	1.00000 551

tuzo >