

PAF 573
Advanced Regression
Two Sample t-tests

```
do "D:\PUAF611\SPRING~1\HANDOUT\JAN28\HAND1.DO"
```

```
. *This program generates tests for crash data in the U.S.;
. *For the year 1996;
.
. #delimit;
delimiter now ;
. *set more 1;
. *First describe data set;
. describe;
```

Contains data from D:\PUAF611\SPRING~1\NHTSAUSA.DTA

```
obs:          44                      NHTSA Crash Data, 1996
vars:         14                      28 Jan 1998 14:20
size:         2,860 (99.7% of memory free)
```

```
-----
1. state      str9    %9s
2. changer    float  %9.0g
3. crashes    float  %9.0g          No. of crashes
4. deaths     float  %9.0g          No. of fatalities
5. dper100p   float  %9.0g          deaths per 100,000 people
6. dper100v   float  %9.0g          deaths per 100,000 vehicles
7. dper100d   float  %9.0g          deaths per 100,000 drivers
8. rural      float  %9.0g          No. of crashes on Rural Highway
9. rper100p   float  %9.0g          rural highway deaths per 100,00
10. rper100v  float  %9.0g          rural highway deaths per 100,00
11. rper100d  float  %9.0g          rural highway deaths per 100,00
12. p         float  %9.0g          Population/100,000
13. v         float  %9.0g          vehicles/100,000
14. d         float  %9.0g          drivers/100,000
-----
```

Sorted by:

```
. *Next summarize data;
. sum;
```

Variable	Obs	Mean	Std. Dev.	Min	Max
state	0				
changer	44	.4772727	.5052578	0	1
crashes	44	731.7727	772.6736	65	3576
deaths	44	825.7727	868.6994	69	3989
dper100p	44	17.02273	5.982489	7	30
dper100v	44	21.84091	7.862017	9	46
dper100d	44	23.65909	8.338599	9	43
rural	44	48.18182	47.30223	0	220
rper100p	44	1.392321	1.384836	0	6.293706
rper100v	44	1.776418	1.719393	0	6.923077
rper100d	44	1.882943	1.738324	0	6.715177
p	44	53.43829	59.69798	4.766667	306.8462
v	44	39.59735	42.0039	4.210526	221.6111

```

d |          44          35.776   38.11626   4.766667   199.45

```

```

. sum if changer==1;

```

Variable	Obs	Mean	Std. Dev.	Min	Max
state	0				
changer	21	1	0	1	1
crashes	21	878.4762	1007.64	80	3576
deaths	21	993	1133.441	85	3989
dper100p	21	20.42857	5.464169	12	30
dper100v	21	26.09524	7.654442	13	46
dper100d	21	28.47619	6.982972	15	43
rural	21	68.33333	57.03888	2	220
rper100p	21	2.094498	1.611023	.3058824	6.293706
rper100v	21	2.653856	1.945898	.3058824	6.923077
rper100d	21	2.840892	1.953046	.3529412	6.715177
p	21	55.16625	73.62132	4.766667	306.8462
v	21	41.0706	52.932	4.333333	221.6111
d	21	36.88934	47.83362	4.766667	199.45

```

. sum if changer==0;

```

Variable	Obs	Mean	Std. Dev.	Min	Max
state	0				
changer	23	0	0	0	0
crashes	23	597.8261	452.301	65	1422
deaths	23	673.087	505.9741	69	1564
dper100p	23	13.91304	4.650641	7	25
dper100v	23	17.95652	5.873502	9	34
dper100d	23	19.26087	7.014384	9	36
rural	23	29.78261	25.96318	0	86
rper100p	23	.751203	.6989706	0	2.925
rper100v	23	.9752792	.9643948	0	4.275
rper100d	23	1.008294	.8733965	0	3.375
p	23	51.86058	45.08758	5.866667	173.7778
v	23	38.25221	29.95677	4.210526	104.2667
d	23	34.75948	27.51993	5.333333	104.2667

```

. *Conduct tests to see if deaths per 100,000 categories are
> *statistically different for changer versus keeper states;
. *Test for deaths per 100,000 people;
. gen dp100pc=dper100p if changer==1;
(23 missing values generated)

. gen dp100pk=dper100p if changer==0;
(21 missing values generated)

```

```
. ttest dp100pc=dp100pk, unpaired;
```

```
Two-sample t test with equal variances      dp100pc: Number of obs =      21
                                             dp100pk: Number of obs =      23
```

Variable	Mean	Std. Err.	t	P> t	[95% Conf. Interval]	
dp100pc	20.42857	1.192379	17.1326	0.0000	17.94131	22.91583
dp100pk	13.91304	.9697256	14.3474	0.0000	11.90196	15.92413
diff	6.515528	1.525533	4.27098	0.0001	3.436877	9.594179

```
Degrees of freedom: 42
```

Ho: mean(dp100pc) - mean(dp100pk) = diff = 0

Ha: diff < 0	Ha: diff ~= 0	Ha: diff > 0
t = 4.2710	t = 4.2710	t = 4.2710
P < t = 0.9999	P > t = 0.0001	P > t = 0.0001

```
. ttest dp100pc=dp100pk, unpaired unequal;
```

```
Two-sample t test with unequal variances    dp100pc: Number of obs =      21
                                             dp100pk: Number of obs =      23
```

Variable	Mean	Std. Err.	t	P> t	[95% Conf. Interval]	
dp100pc	20.42857	1.192379	17.1326	0.0000	17.94131	22.91583
dp100pk	13.91304	.9697256	14.3474	0.0000	11.90196	15.92413
diff	6.515528	1.536924	4.23933	0.0001	3.408057	9.622999

```
Satterthwaite's degrees of freedom: 39.49764
```

Ho: mean(dp100pc) - mean(dp100pk) = diff = 0

Ha: diff < 0	Ha: diff ~= 0	Ha: diff > 0
t = 4.2393	t = 4.2393	t = 4.2393
P < t = 0.9999	P > t = 0.0001	P > t = 0.0001

```
. *Test for deaths per 100,000 vehicles;
```

```
. gen dp100vc=dper100v if changer==1;
(23 missing values generated)
```

```
. gen dp100vk=dper100v if changer==0;
(21 missing values generated)
```

```
. ttest dp100vc=dp100vk, unpaired;
```

```
Two-sample t test with equal variances      dp100vc: Number of obs =      21
                                             dp100vk: Number of obs =      23
```

Variable	Mean	Std. Err.	t	P> t	[95% Conf. Interval]	
dp100vc	26.09524	1.670336	15.6227	0.0000	22.61098	29.5795
dp100vk	17.95652	1.22471	14.6619	0.0000	15.41663	20.49641
diff	8.138716	2.046411	3.97707	0.0003	4.008891	12.26854

```
Degrees of freedom: 42
```

Ho: mean(dp100vc) - mean(dp100vk) = diff = 0

Ha: diff < 0	Ha: diff ~= 0	Ha: diff > 0
t = 3.9771	t = 3.9771	t = 3.9771
P < t = 0.9999	P > t = 0.0003	P > t = 0.0001

```
. ttest dp100vc=dp100vk, unpaired unequal;
```

```
Two-sample t test with unequal variances    dp100vc: Number of obs =      21
                                             dp100vk: Number of obs =      23
```

Variable	Mean	Std. Err.	t	P> t	[95% Conf. Interval]	
dp100vc	26.09524	1.670336	15.6227	0.0000	22.61098	29.5795
dp100vk	17.95652	1.22471	14.6619	0.0000	15.41663	20.49641
diff	8.138716	2.071216	3.92944	0.0004	3.943719	12.33371

```
Satterthwaite's degrees of freedom: 37.445764
```

Ho: mean(dp100vc) - mean(dp100vk) = diff = 0

Ha: diff < 0	Ha: diff ~= 0	Ha: diff > 0
t = 3.9294	t = 3.9294	t = 3.9294
P < t = 0.9998	P > t = 0.0004	P > t = 0.0002

```
. *Test for deaths per 100,000 drivers;
```

```
. gen dp100dc=dper100d if changer==1;
(23 missing values generated)
```

```
. gen dp100dk=dper100d if changer==0;
(21 missing values generated)
```

```
. ttest dp100dc=dp100dk, unpaired;
```

```
Two-sample t test with equal variances      dp100dc: Number of obs =      21
                                             dp100dk: Number of obs =      23
```

Variable	Mean	Std. Err.	t	P> t	[95% Conf. Interval]	
dp100dc	28.47619	1.52381	18.6875	0.0000	25.29758	31.6548
dp100dk	19.26087	1.4626	13.1689	0.0000	16.22762	22.29412
diff	9.215321	2.112595	4.36209	0.0001	4.951931	13.47871

Degrees of freedom: 42

Ho: mean(dp100dc) - mean(dp100dk) = diff = 0

Ha: diff < 0	Ha: diff ~= 0	Ha: diff > 0
t = 4.3621	t = 4.3621	t = 4.3621
P < t = 1.0000	P > t = 0.0001	P > t = 0.0000

```
. ttest dp100dc=dp100dk, unpaired unequal;
```

```
Two-sample t test with unequal variances    dp100dc: Number of obs =      21
                                             dp100dk: Number of obs =      23
```

Variable	Mean	Std. Err.	t	P> t	[95% Conf. Interval]	
dp100dc	28.47619	1.52381	18.6875	0.0000	25.29758	31.6548
dp100dk	19.26087	1.4626	13.1689	0.0000	16.22762	22.29412
diff	9.215321	2.112154	4.363	0.0001	4.951828	13.47881

Satterthwaite's degrees of freedom: 41.672178

Ho: mean(dp100dc) - mean(dp100dk) = diff = 0

Ha: diff < 0	Ha: diff ~= 0	Ha: diff > 0
t = 4.3630	t = 4.3630	t = 4.3630
P < t = 1.0000	P > t = 0.0001	P > t = 0.0000

```
. *Conduct tests for rural highways only;
```

```
. *Test for deaths per 100,000 people;
```

```
. gen rp100pc=rper100p if changer==1;
```

```
(23 missing values generated)
```

```
. gen rp100pk=rper100p if changer==0;
```

```
(21 missing values generated)
```

```
. ttest rp100pc=rp100pk, unpaired;
```

```
Two-sample t test with equal variances      rp100pc: Number of obs =      21
                                             rp100pk: Number of obs =      23
```

Variable	Mean	Std. Err.	t	P> t	[95% Conf. Interval]	
rp100pc	2.094498	.3515541	5.95783	0.0000	1.361169	2.827827
rp100pk	.751203	.1457454	5.15421	0.0000	.4489455	1.053461
diff	1.343295	.3686467	3.64386	0.0007	.5993362	2.087254

Degrees of freedom: 42

Ho: mean(rp100pc) - mean(rp100pk) = diff = 0

Ha: diff < 0	Ha: diff ~= 0	Ha: diff > 0
t = 3.6439	t = 3.6439	t = 3.6439
P < t = 0.9996	P > t = 0.0007	P > t = 0.0004

```
. ttest rp100pc=rp100pk, unpaired unequal;
```

```
Two-sample t test with unequal variances    rp100pc: Number of obs =      21
                                             rp100pk: Number of obs =      23
```

Variable	Mean	Std. Err.	t	P> t	[95% Conf. Interval]	
rp100pc	2.094498	.3515541	5.95783	0.0000	1.361169	2.827827
rp100pk	.751203	.1457454	5.15421	0.0000	.4489455	1.053461
diff	1.343295	.3805681	3.52971	0.0015	.5620889	2.124502

Satterthwaite's degrees of freedom: 26.747396

Ho: mean(rp100pc) - mean(rp100pk) = diff = 0

Ha: diff < 0	Ha: diff ~= 0	Ha: diff > 0
t = 3.5297	t = 3.5297	t = 3.5297
P < t = 0.9992	P > t = 0.0015	P > t = 0.0008

```
. *Test for deaths per 100,000 vehicles;
```

```
. gen rp100vc=rper100v if changer==1;
(23 missing values generated)
```

```
. gen rp100vk=rper100v if changer==0;
(21 missing values generated)
```

```
. ttest rp100vc=rp100vk, unpaired;
```

```
Two-sample t test with equal variances      rp100vc: Number of obs =      21
                                             rp100vk: Number of obs =      23
```

Variable	Mean	Std. Err.	t	P> t	[95% Conf. Interval]	
rp100vc	2.653856	.4246299	6.24981	0.0000	1.768094	3.539618
rp100vk	.9752792	.2010902	4.84996	0.0001	.5582436	1.392315
diff	1.678577	.4567693	3.67489	0.0007	.7567791	2.600375

Degrees of freedom: 42

Ho: mean(rp100vc) - mean(rp100vk) = diff = 0

Ha: diff < 0	Ha: diff ~= 0	Ha: diff > 0
t = 3.6749	t = 3.6749	t = 3.6749
P < t = 0.9997	P > t = 0.0007	P > t = 0.0003

```
. ttest rp100vc=rp100vk, unpaired unequal;
```

```
Two-sample t test with unequal variances    rp100vc: Number of obs =      21
                                             rp100vk: Number of obs =      23
```

Variable	Mean	Std. Err.	t	P> t	[95% Conf. Interval]	
rp100vc	2.653856	.4246299	6.24981	0.0000	1.768094	3.539618
rp100vk	.9752792	.2010902	4.84996	0.0001	.5582436	1.392315
diff	1.678577	.4698381	3.57267	0.0013	.7171633	2.63999

Satterthwaite's degrees of freedom: 28.66581

Ho: mean(rp100vc) - mean(rp100vk) = diff = 0

Ha: diff < 0	Ha: diff ~= 0	Ha: diff > 0
t = 3.5727	t = 3.5727	t = 3.5727
P < t = 0.9994	P > t = 0.0013	P > t = 0.0006

```
. *Test for deaths per 100,000 drivers;
```

```
. gen rp100dc=rper100d if changer==1;
(23 missing values generated)
```

```
. gen rp100dk=rper100d if changer==0;
(21 missing values generated)
```

```
. ttest rp100dc=rp100dk, unpaired;
```

```
Two-sample t test with equal variances      rp100dc: Number of obs =      21
                                             rp100dk: Number of obs =      23
```

Variable	Mean	Std. Err.	t	P> t	[95% Conf. Interval]	
rp100dc	2.840892	.4261895	6.6658	0.0000	1.951876	3.729908
rp100dk	1.008294	.1821158	5.53656	0.0000	.6306093	1.385979
diff	1.832598	.449296	4.07882	0.0002	.9258819	2.739314

Degrees of freedom: 42

Ho: mean(rp100dc) - mean(rp100dk) = diff = 0

Ha: diff < 0	Ha: diff ~= 0	Ha: diff > 0
t = 4.0788	t = 4.0788	t = 4.0788
P < t = 0.9999	P > t = 0.0002	P > t = 0.0001

```
. ttest rp100dc=rp100dk, unpaired unequal;
```

```
Two-sample t test with unequal variances    rp100dc: Number of obs =      21
                                             rp100dk: Number of obs =      23
```

Variable	Mean	Std. Err.	t	P> t	[95% Conf. Interval]	
rp100dc	2.840892	.4261895	6.6658	0.0000	1.951876	3.729908
rp100dk	1.008294	.1821158	5.53656	0.0000	.6306093	1.385979
diff	1.832598	.4634692	3.95409	0.0005	.88188	2.783316

Satterthwaite's degrees of freedom: 27.147782

Ho: mean(rp100dc) - mean(rp100dk) = diff = 0

Ha: diff < 0	Ha: diff ~= 0	Ha: diff > 0
t = 3.9541	t = 3.9541	t = 3.9541
P < t = 0.9998	P > t = 0.0005	P > t = 0.0002

```
. *end program;
```

```
.  
end of do-file
```

```
. exit, clear  
exit, clear
```