Hypothesis test in regression is ①

$H_0: \beta_1 = 0$ (β_1 is the slope)

$H_1: \beta_1 \neq 0$

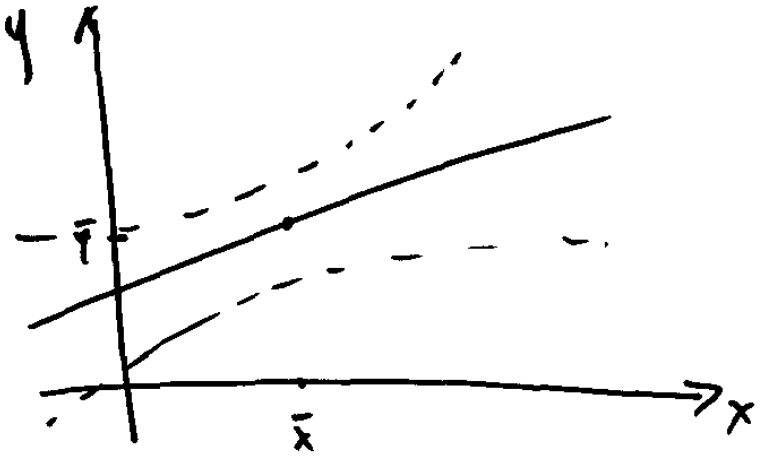
H_0: "x is useless as a linear predictor of y"

From the handout, $\hat{y} = -240.5 + 5.5x$

$r^2 = .922$

p-val = .049 so reject H_0. At $\alpha = .05$

②



Multiple regression model (2nd handout)

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \varepsilon$$

$H_0: \beta_1 = \beta_2 = \beta_3 = \ldots = 0$

Why was the F test significant, but the t-test for each slope was insignificant?

Answer: Our 3 predictors are correlated with each other.

"multicollinearity"

Homework assignment on following pages

11. Worried about Retirement? In February 2008, the Gallup organization surveyed 1,034 adults between 30 and 64 years of age and found that 548 were worried that they will outlive their money after they retire. Does the sample evidence suggest that a majority of 30- to 64-year-olds in the United States are worried they will outlive their money? Use the $\alpha = 0.05$ level of significance.

17. Abstain from Alcohol In November 1945, the Gallup organization surveyed 1,100 adult Americans and asked, "Are you a total abstainer from, or do you on occasion consume, alcoholic beverages?" Of the 1,100 adults surveyed, 363 indicated that they were total abstainers. In July 2007, the same question was asked of 1,100 adult Americans and 396 indicated that they were total abstainers. Has the proportion of adult Americans who totally abstain from alcohol changed? Use the $\alpha = 0.05$ level of significance.

16. Rat's Hemoglobin Hemoglobin helps the red blood cells transport oxygen and remove carbon dioxide. Researchers at NASA wanted to determine the effects of space flight on a rat's hemoglobin. The following data represent the hemoglobin (in grams per deciliter) at lift-off minus 3 days (H-L3) and immediately upon the return (H-R0) for 12 randomly selected rats sent to space on the Spacelab Sciences 1 flight.

Rat	1	2	3	4	5	6
H-L3, X_i	15.2	16.1	15.3	16.4	15.7	14.7
H-R0, Y_i	15.8	16.5	16.7	15.7	16.9	13.1
Rat	**7**	**8**	**9**	**10**	**11**	**12**
H-L3, X_i	14.3	14.5	15.2	16.1	15.1	15.8
H-R0, Y_i	16.4	16.5	16.0	16.8	17.6	16.9

Source: NASA Life Sciences Data Archive

Does the evidence suggest that the hemoglobin levels at lift-off minus 3 days are less than the hemoglobin levels upon return at the $\alpha = 0.05$ level of significance?

Note: A normal probability plot and boxplot of the data indicate that the differences are approximately normally distributed with no outliers.

13. Concrete Strength An engineer wanted to know whether the strength of two different concrete mix designs differed significantly. He randomly selected 9 cylinders, measuring 6 inches in diameter and 12 inches in height, into which mixture 67-0-301 was poured. After 28 days, he measured the strength (in pounds per square inch) of the cylinder. He also randomly selected 10 cylinders of mixture 67-0-400 and performed the same test. The results are as follows:

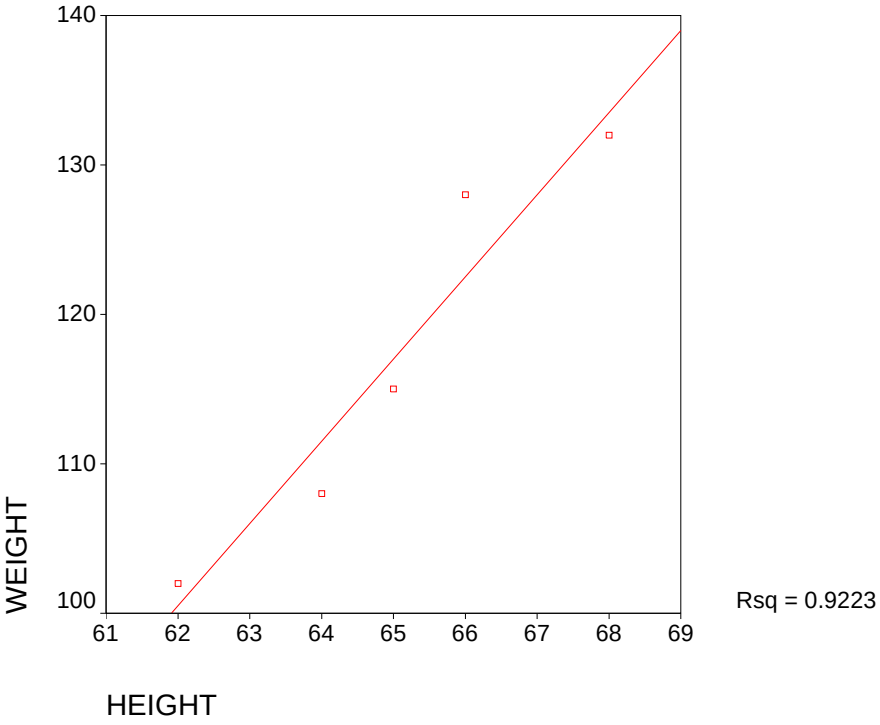
Mixture 67-0-301			Mixture 67-0-400			
3,960	4,090	3,100	4,070	4,890	5,020	4,330
3,830	3,200	3,780	4,640	5,220	4,190	3,730
4,080	4,040	2,940	4,120	4,620		

Note: Normal probability plots indicate that the data are approximately normal and boxplots indicate that there are no outliers.

Determine whether mixture 67-0-400 is stronger than mixture 67-0-301 at the $\alpha = 0.05$ level of significance.

	height	weight
1	68	132
2	64	108
3	62	102
4	65	115
5	66	128
6	63	.

Graph



Regression

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	HEIGHT[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: WEIGHT

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.960[a]	.922	.896	4.123

a. Predictors: (Constant), HEIGHT

b. Dependent Variable: WEIGHT

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	605.000	1	605.000	35.588	.009[a]
	Residual	51.000	3	17.000		
	Total	656.000	4			

a. Predictors: (Constant), HEIGHT

b. Dependent Variable: WEIGHT

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-240.500	59.955		-4.011	.028
	HEIGHT	5.500	.922	.960	5.966	.009

a. Dependent Variable: WEIGHT

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	100.50	133.50	117.00	12.298	5
Std. Predicted Value	-1.342	1.342	.000	1.000	5
Standard Error of Predicted Value	1.844	3.324	2.523	.737	5
Adjusted Predicted Value	97.71	136.29	116.97	13.931	5
Residual	-3.50	5.50	.00	3.571	5
Std. Residual	-.849	1.334	.000	.866	5
Stud. Residual	-.980	1.540	.004	1.047	5
Deleted Residual	-4.67	7.33	.03	5.444	5
Stud. Deleted Residual	-.971	2.750	.263	1.496	5
Mahal. Distance	.000	1.800	.800	.917	5
Cook's Distance	.037	.395	.259	.154	5
Centered Leverage Value	.000	.450	.200	.229	5

a. Dependent Variable: WEIGHT

	squarefo	bedrooms	baths	askingpr
1	3632	4	2.5	419
2	4889	6	5.0	399
3	3000	5	3.5	395
4	3669	4	3.5	379
5	2800	4	3.0	359
6	3600	5	3.5	349
7	2800	5	2.5	320
8	2257	3	3.0	299
9	2000	3	3.0	295
10	2455	3	2.5	290
11	2250	3	2.0	285
12	2938	3	2.0	269
13	2399	3	2.0	260

Correlations

Correlations

		SQUAREFO	BEDROOMS	BATHS	ASKINGPR
SQUAREFO	Pearson Correlation	1	.791**	.731**	.749**
	Sig. (2-tailed)	.	.001	.005	.003
	N	13	13	13	13
BEDROOMS	Pearson Correlation	.791**	1	.759**	.738**
	Sig. (2-tailed)	.001	.	.003	.004
	N	13	13	13	13
BATHS	Pearson Correlation	.731**	.759**	1	.669*
	Sig. (2-tailed)	.005	.003	.	.012
	N	13	13	13	13
ASKINGPR	Pearson Correlation	.749**	.738**	.669*	1
	Sig. (2-tailed)	.003	.004	.012	.
	N	13	13	13	13

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Regression

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	BATHS, SQUAREFO, BEDROOMS[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: ASKINGPR

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.790[a]	.625	.500	38.191

a. Predictors: (Constant), BATHS, SQUAREFO, BEDROOMS

b. Dependent Variable: ASKINGPR

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	21854.557	3	7284.852	4.995	.026[a]
	Residual	13127.135	9	1458.571		
	Total	34981.692	12			

a. Predictors: (Constant), BATHS, SQUAREFO, BEDROOMS

b. Dependent Variable: ASKINGPR

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	161.168	45.497		3.542	.006
	SQUAREFO	.027	.024	.396	1.120	.292
	BEDROOMS	16.759	19.285	.322	.869	.407
	BATHS	8.699	21.402	.135	.406	.694

Coefficients[a]

Model		Collinearity Statistics	
		Tolerance	VIF
1	(Constant)		
	SQUAREFO	.334	2.996
	BEDROOMS	.303	3.295
	BATHS	.378	2.645

a. Dependent Variable: ASKINGPR

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	SQUAREFO	BEDROOMS	BATHS
1	1	3.928	1.000	.00	.00	.00	.00
	2	.042	9.725	.97	.03	.03	.09
	3	.018	14.646	.02	.37	.08	.86
	4	.013	17.678	.00	.60	.89	.06

a. Dependent Variable: ASKINGPR

Residuals Statistics[a]

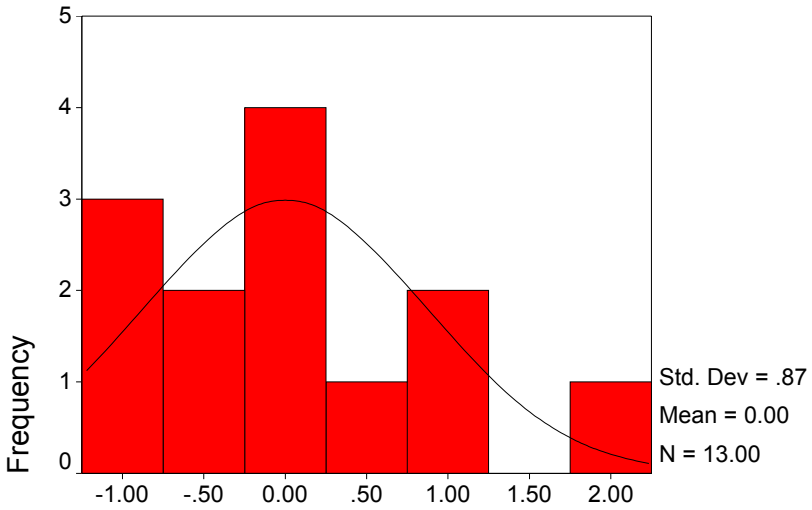
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	289.18	436.33	332.15	42.676	13
Std. Predicted Value	-1.007	2.441	.000	1.000	13
Standard Error of Predicted Value	11.998	31.019	20.478	5.648	13
Adjusted Predicted Value	288.61	508.69	337.34	59.477	13
Residual	-38.63	71.65	.00	33.075	13
Std. Residual	-1.012	1.876	.000	.866	13
Stud. Residual	-1.675	2.342	-.046	1.107	13
Deleted Residual	-109.69	111.67	-5.19	56.455	13
Stud. Deleted Residual	-1.904	3.533	.029	1.373	13
Mahal. Distance	.261	6.993	2.769	2.034	13
Cook's Distance	.000	1.360	.226	.402	13
Centered Leverage Value	.022	.583	.231	.170	13

a. Dependent Variable: ASKINGPR

Charts

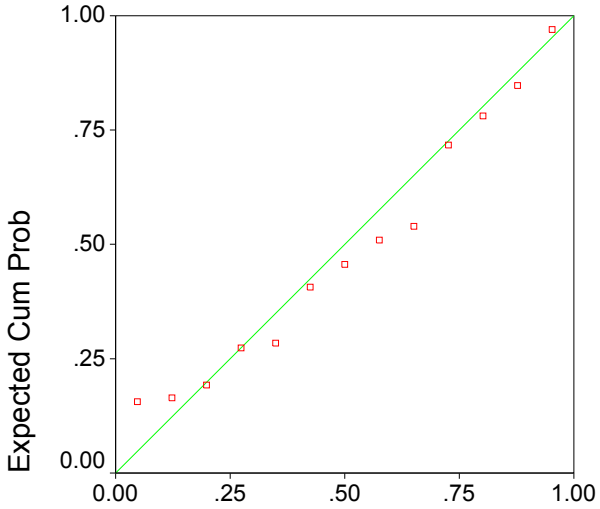
Histogram

Dependent Variable: ASKINGPR



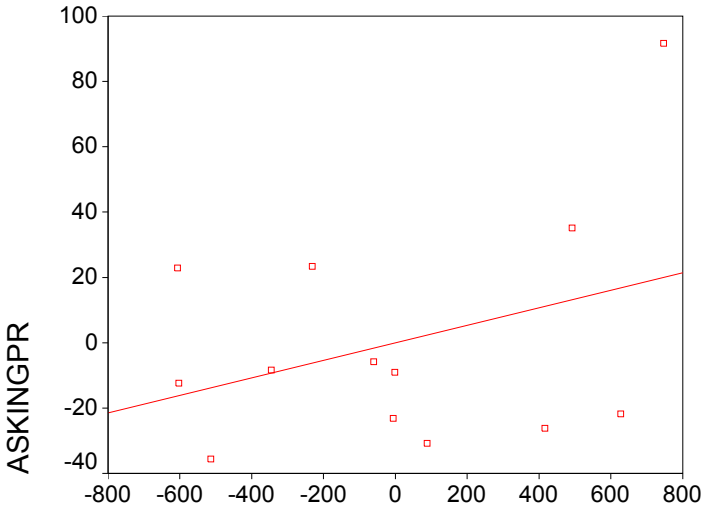
Normal P-P Plot of Regression Sta

Dependent Variable: ASKINGPR



Partial Regression Plot

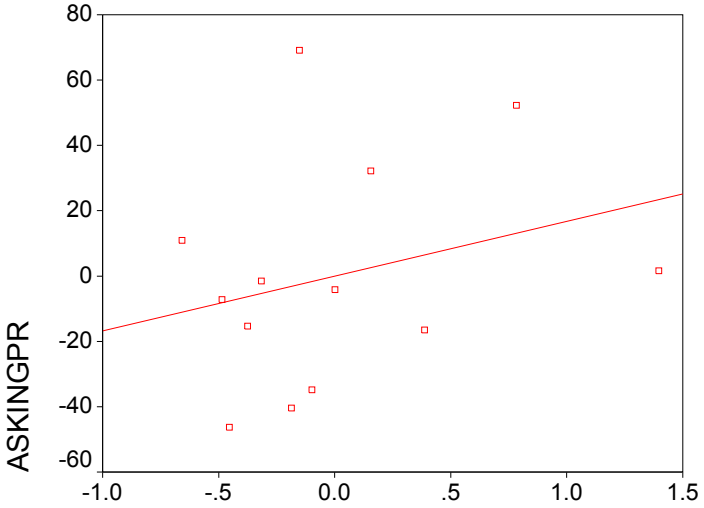
Dependent Variable: ASKINGPR



Rsq = 0.1224

Partial Regression Plot

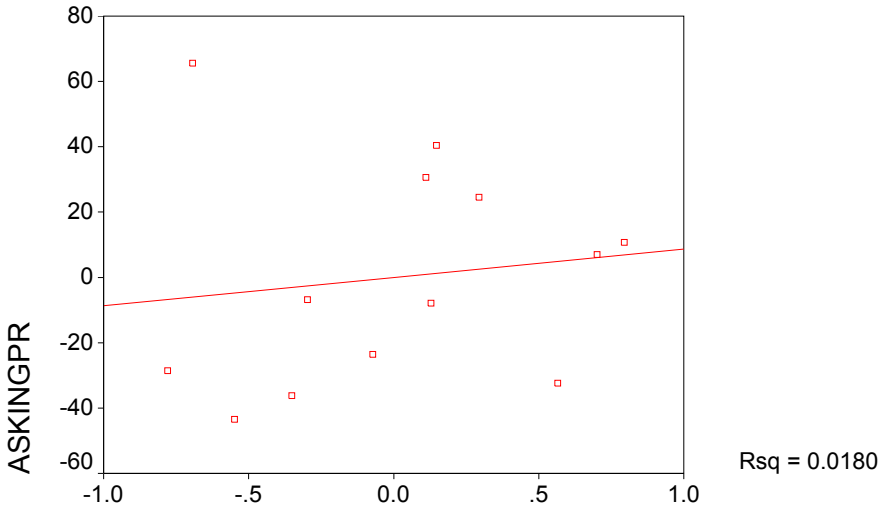
Dependent Variable: ASKINGPR



Rsq = 0.0774

Partial Regression Plot

Dependent Variable: ASKINGPR



Regression

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	SQUAREFO	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

a. Dependent Variable: ASKINGPR

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.749[a]	.562	.522	37.333

a. Predictors: (Constant), SQUAREFO

b. Dependent Variable: ASKINGPR

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	19650.762	1	19650.762	14.099	.003[a]
	Residual	15330.931	11	1393.721		
	Total	34981.692	12			

a. Predictors: (Constant), SQUAREFO

b. Dependent Variable: ASKINGPR

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	181.066	41.548		4.358	.001
	SQUAREFO	.051	.014	.749	3.755	.003

Coefficients[a]

Model		Collinearity Statistics	
		Tolerance	VIF
1	(Constant)		
	SQUAREFO	1.000	1.000

a. Dependent Variable: ASKINGPR

Excluded Variables[b]

Model		Beta In	t	Sig.	Partial Correlation
1	BEDROOMS	.387[a]	1.212	.253	.358
	BATHS	.260[a]	.880	.399	.268

Excluded Variables[b]

Model		Collinearity Statistics		
		Tolerance	VIF	Minimum Tolerance
1	BEDROOMS	.374	2.675	.374
	BATHS	.466	2.148	.466

a. Predictors in the Model: (Constant), SQUAREFO

b. Dependent Variable: ASKINGPR

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions	
				(Constant)	SQUAREFO
1	1	1.968	1.000	.02	.02
	2	.032	7.899	.98	.98

a. Dependent Variable: ASKINGPR

Residuals Statistics[a]

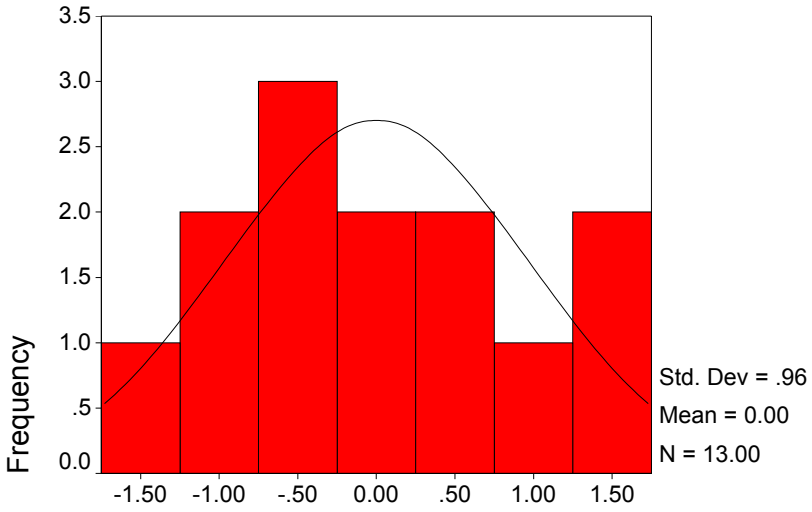
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	282.60	429.27	332.15	40.467	13
Std. Predicted Value	-1.225	2.400	.000	1.000	13
Standard Error of Predicted Value	10.359	27.859	13.963	4.591	13
Adjusted Predicted Value	279.46	467.30	334.70	48.165	13
Residual	-61.22	61.63	.00	35.743	13
Std. Residual	-1.640	1.651	.000	.957	13
Stud. Residual	-1.707	1.718	-.026	1.039	13
Deleted Residual	-68.30	66.77	-2.55	43.156	13
Stud. Deleted Residual	-1.898	1.916	-.020	1.112	13
Mahal. Distance	.001	5.759	.923	1.515	13
Cook's Distance	.000	.932	.121	.251	13
Centered Leverage Value	.000	.480	.077	.126	13

a. Dependent Variable: ASKINGPR

Charts

Histogram

Dependent Variable: ASKINGPR



Normal P-P Plot of Regression Sta

Dependent Variable: ASKINGPR