Freund & Wilson (1997) : Prediction of weight of wood from trees (Table 8.24)

Observation	Dbh	Weight	Dbh*Dbh	Wt*Wt	Dbh*Wt	Predicted	Residual
1	5.7	174	32.49	30276	991.8	288.42	-114.42
2	8.1	745	65.61	555025	6034.5	716.97	28.03
3	8.3	814	68.89	662596	6756.2	752.68	61.32
4	7.0	408	49.00	166464	2856.0	520.55	-112.55
5	6.2	226	38.44	51076	1401.2	377.7	-151.7
6	11.4	1675	129.96	2805625	19095.0	1306.23	368.77
7	11.6	1491	134.56	2223081	17295.6	1341.94	149.06
8	4.5	121	20.25	14641	544.5	74.14	46.86
9	3.5	58	12.25	3364	203.0	-104.42	162.42
10	6.2	278	38.44	77284	1723.6	377.7	-99.7
11	5.7	220	32.49	48400	1254.0	288.42	-68.42
12	6.0	342	36.00	116964	2052.0	341.99	0.01
13	5.6	209	31.36	43681	1170.4	270.56	-61.56
14	4.0	84	16.00	7056	336.0	-15.14	99.14
15	6.7	313	44.89	97969	2097.1	466.98	-153.98
16	4.0	60	16.00	3600	240.0	-15.14	75.14
17	12.1	1692	146.41	2862864	20473.2	1431.22	260.78
18	4.5	74	20.25	5476	333.0	74.14	-0.14
19	8.6	515	73.96	265225	4429.0	806.25	-291.25
20	9.3	766	86.49	586756	7123.8	931.25	-165.25
21	6.5	345	42.25	119025	2242.5	431.27	-86.27
22	5.6	210	31.36	44100	1176.0	270.56	-60.56
23	4.3	100	18.49	10000	430.0	38.43	61.57
24	4.5	122	20.25	14884	549.0	74.14	47.86
25	7.7	539	59.29	290521	4150.3	645.54	-106.54
26	8.8	815	77.44	664225	7172.0	841.96	-26.96
27	5.0	194	25.00	37636	970.0	163.42	30.58
28	5.4	280	29.16	78400	1512.0	234.85	45.15
29	6.0	296	36.00	87616	1776.0	341.99	-45.99
30	7.4	462	54.76	213444	3418.8	591.98	-129.98
31	5.6	200	31.36	40000	1120.0	270.56	-70.56
32	5.5	229	30.25	52441	1259.5	252.7	-23.7
33	4.3	125	18.49	15625	537.5	38.43	86.57
34	4.2	84	17.64	7056	352.8	20.57	63.43
35	3.7	70	13.69	4900	259.0	-68.71	138.71
36	6.1	224	37.21	50176	1366.4	359.84	-135.84
37	3.9	99	15.21	9801	386.1	-33	132
38	5.2	200	27.04	40000	1040.0	199.14	0.86
39	5.6	214	31.36	45796	1198.4	270.56	-56.56
40	7.8	712	60.84	506944	5553.6	663.4	48.6
41	6.1	297	37.21	88209	1811.7	359.84	-62.84
42	6.1	238	37.21	56644	1451.8	359.84	-121.84
43	4.0	89	16.00	7921	356.0	-15.14	104.14
44	4.0	76	16.00	5776	304.0	-15.14	91.14
45	8.0	614	64.00	376996	4912.0	699.11	-85.11
46	5.2	194	27.04	37636	1008.8	199.14	-5.14
47	3.7	66	13.69	4356	244.2	-68.71	134.71
Sum	**289.2**	**17359**	**1981.98**	**13537551**	**142968.3**	**Sum =**	**0**
Mean	**6.15**	**369.34**	**42.17**	**288033**	**3041.9**	**SS =**	**670190.732**
n	**47**	**47**	**47**	**47**	**47**		

Intermediate Calculations

Sum X =	289.2	Sum Y =	17359
Sum X^2 =	1981.98	Sum Y^2 =	13537551
Mean X=	6.153191489	Mean Y=	369.3404255
Sum XY =	142968.3	n =	47

Correction factors and Corrected values (Sums of squares and cross-products)

CF for X =	C_{xx} =	1779.502979	Corrected SS X =	S_{xx} =202.4770213
CF for Y =	C_{yy} =	6411380.447	Corrected SS Y =	S_{yy} =7126170.553
CF for XY =	C_{xy} =	106813.2511	Corrected SS XY =	S_{xy} =36155.04894

Model Parameter Estimates

Slope = b_1 = 36155.04894 / 202.4770213 =178.5637141
Intercept = b_0 = 369.3404255 - 178.5637141 * 6.153191489 = -729.3963003
Regression Line $Y_i = b_0 + b_1 * X_i + e_i$
$Y_i = -729.3963003 + 178.5637141 * X_i + e_i$

ANOVA Table

SSTotal =7126170.553
SSRegression 36155.04894^2 / 202.4770213 = 6455979.821
SSError = 7126170.553 - 6455979.821 = 670190.7322

Source	**df**	**SS**	**MS**	**F**
Regression	1	6455979.821	6455979.821	433.4871821
Error	45	670190.7322	14893.12738	
Total	46	7126170.553		

Standard error of b_1 : where $t_{(0.05/2, 45 df)}$ = 2.014103 ; $S_{b_1} = \sqrt{MSE / \sum X_i^2}$ = 8.576401034

P(178.5637 - 2.0141*8.5764 ≤ β_1 ≤ 178.5637 + 2.0141*8.5764) = 0.95

P(161.289956 ≤ β_1 ≤ 195.8375) = 0.95

Testing b_1 against a specified value : H_0: β_1 = 200 versus H_1: $\beta_1 \neq 200$

$t = \frac{b_1 - \beta_{1|H_0}}{S_{b_1}}$ = (178.5637141 - 200) / 8.576401034 = -2.49945

Note that t^2 = F = 6.247251 ; This test would be done in SAS as an F statement

The variance of a linear combination is given by the sum of the variances plus twice the covariances.

e.g. for **A = aX + bY + cZ**

then **Var(A) =** $a^2\sigma^2_X + b^2\sigma^2_Y + c^2\sigma^2_Z + 2(ab\sigma_{XY} + ac\sigma_{XZ} + bc\sigma_{YZ})$

where the covariances are equal to zero if the variables are independent

For the linear combination $\hat{Y}_i = b_0 + b_1 X_i$ **, the standard error of** $\hat{Y}_i$ **is as follows.**

Standard error of the regression line ($\hat{Y}_i$**):** $S_{\hat{Y}_{y|x}} = \sqrt{MSE\left(\frac{1}{n} + \frac{(X_i - \bar{X}.)^2}{\sum (X_i - \bar{X}.)^2}\right)}$

The calculation above DOES NOT assume that the covariances of the regression coefficients are independent. However, for the variance of individual points the linear combination is $Y_i = b_0 + b_1 X_i + e_i = \hat{Y}_i + e_i$**. For this linear combination the terms for the predicted value and residuals are assumed independent (i.e.** $\hat{Y}_i$ **is independent of** e_i**).**

$$S_{\hat{Y}_{y|x}} = \sqrt{MSE\left(\frac{1}{n} + \frac{(X_i - \bar{X}.)^2}{S_{xx}}\right) + MSE} = \sqrt{MSE\left(1 + \frac{1}{n} + \frac{(X_i - \bar{X}.)^2}{S_{xx}}\right)}$$

Standard error of b_0 **is the same as the standard error of the regression line where** $X_i = 0$

SQRT(14893.12738(0.021276596 + (0 - 37.8617655) / 202.4770213 = 55.69366336

Confidence interval on b_0 where b_0 = -729.3963003 and $t_{(0.05/2,\ 45\ df)}$ = 2.014103

P(-729.3963 - 2.0141*55.6937 ≤ β_0 ≤ -729.3963+2.0141*55.6937)=0.95

P(-841.5690916 ≤ β_0 ≤ -617.223509) = 0.95

Estimate and standard error of an individual observation (e.g. the weight of wood for a ten-inch-diameter tree)

Y =-729.3963003 + 178.5637141* X = -729.3963003 + 178.5637141 * 10 =1056.240841

se($b_{x=10}$) =14893.1274*(1 + 0.02128 + (10 - 14.79794) / 202.4770) = 127.6654

P(1056.2408-2.0141*127.6654 ≤ $\mu_{x=10}$ ≤ 1056.2408+2.0141*127.6654)=0.95

P(799.1094964 ≤ $\mu_{x=10}$ ≤ 1313.372185) = 0.95

Calculate the coefficient of Determination and correlation

R^2 = 0.905953594 or 90.59535936 %

r = 0.951815945