4.11 Two quality control technicians measured the surface finish of a metal part, obtaining the data in Table 4E.1. Assume that the measurements are normally distributed.

Technician 1	Technician 2	x =	x =
1.45	1.54	1.383	1.376
1.37	1.41	s =	s =
1.21	1.56	0.005	0.027
1.54	1.37	0.000	0.001
1.48	1.2	0.030	0.034
1.29	1.31	0.025	0.000
1.34	1.27	0.009	0.031
	1.35	0.009	0.004
		0.002	0.011
n = 7	n = 8		0.001
		---------------	---------------
		0.115	0.125
		Spool =	
		0.1204	

(a) Test the hypothesis that the mean surface finish measurements made by the two technicians are equal. Use α = 0.05, and assume equal variances.

$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 \neq \mu_2$

$\sigma_1 \sigma_2$ 未知但相等

$t_{0.025}(13) = 2.166$

$t = (\bar{x}_1 - \bar{x}_2 / spool(\sqrt{}(1/n_1 + 1/n_2))) = 0.106 < 2.166$

reject H_0

(b) What are the practical implications of the test in part (a)? Discuss what practical conclusions you would draw if the null hypothesis were rejected.

The mean surface finish measurements made by two technicians are not statistically different at α = 0.05

(c) Assuming that the variances are equal, construct a 95% confidence interval on the mean difference in surface-finish measurements.

$\bar{x}_1 - \bar{x}_2 \pm t^*spool(\sqrt{}(1/n_1 + 1/n_2))$

$1.383-1.376 \pm 2.16(0.1204)(\sqrt{}(1/n_1 + 1/n_2))$

$= (-0.128, 0.141)$

(d) Test the hypothesis that the variances of the measurements made by the two technicians are equal. Use α = 0.05. What are the practical implications if the null hypothesis is rejected?

$H_0 : \sigma_1^2 = \sigma_2^2 \quad H_1 : \sigma_1^2 \neq \sigma_2^2$

$F = s_1^2 / s_2^2 = 0.85$

p-value = 0.57 > 0.05

We fail to reject H_0, there is no difference the variance of mean surface finish measurement between the two techniques.

(e) Construct a 95% confidence interval estimate of the ratio of the variances of technician measurement error.

$[(s_1^2/s_2^2)/(F_{0.05/2,6,7}), (s_1^2/s_2^2)/(F_{(1-0.05)/2,6,7})]$

$= [(0.115^2/0.125^2)/5.1186, (0.115^2/0.125^2)/0.1756]$

$= [0.1654, 4.8206]$

(f) Construct a 95% confidence interval on the variance of measurement error for technician 2.

$[(n_2-1)S_2^2/x^2_{0.05/2}(n_2-1), (n_2-1)S_2^2/x^2_{1-0.05/2}(n_2-1)]$

$= [7 * 0.125^2/16.0128, 7 * 0.125^2/1.6899]$

$=[0.0068, 0.0697]$

(g) Does the normality assumption seem reasonable for the data?

7	Number of data points	S	-49.99581
1.383	Sample Mean		
0.115	Sample Sigma	p-value calculations	
		p1	0
0.1423	AD test statistic	p2	0
0.164033	AD* test statistic	p3	0
0.943049	P-value	p4	0.9430494

$$AD = -N - \frac{2i-1}{N}(\ln(F(Y_i)) + \ln(1-F(Y_{N+1-i})))$$

$$AD^* = AD\left(1 + \frac{0.75}{N} + \frac{2.25}{N^2}\right)$$

ENTER DATA HERE	Sorted	Count	F1i	1-F1i	F2i	Si	N Plt Line	N Plt Line
1.45	1.21	1	0.0661529	0.9338471	0.0856174	-5.173653	1.226146	0.086207
1.37	1.29	2	0.2093987	0.7906013	0.1988251	-9.536534	1.295767	0.224138
1.21	1.34	3	0.3545154	0.6454846	0.2794035	-11.56051	1.342323	0.362069
1.54	1.37	4	0.4554325	0.5445675	0.5445675	-9.759898	1.382857	0.5
1.48	1.45	5	0.7205965	0.2794035	0.6454846	-6.888868	1.423392	0.637931
1.29	1.48	6	0.8011749	0.1988251	0.7906013	-5.023013	1.469947	0.775862
1.34	1.54	7	0.9143826	0.0856174	0.9338471	-2.053334	1.539568	0.913793

Technician 1

Test Hypotheses

H_0: Data is sampled from a population that is normally distributed (no difference between the data and normal data).

H_A: Data is sampled from a population that is not normally distributed.

		Count OK?	Seems OK...
8	Number of data points	S	-65.880688
1.376	Sample Mean		
0.125	Sample Sigma	p-value calculations	
		p1	0
0.2351	AD test statistic	p2	0
0.26539	AD* test statistic	p3	0.6933086
0.693309	P-value	p4	0

$$AD = -N - \frac{2i-1}{N}(\ln(F(Y_i)) + \ln(1-F(Y_{N+1-i})))$$

$$AD^* = AD\left(1 + \frac{0.75}{N} + \frac{2.25}{N^2}\right)$$

ENTER DATA HERE	Sorted	Count	F1i	1-F1i	F2i	Si	N Plt Line	N Plt Line
1.54	1.2	1	0.07909133	0.92090867	0.0706102	-5.1877328	1.197129	0.075758
1.41	1.27	2	0.19745981	0.80254019	0.0949079	-11.931206	1.269779	0.19697
1.56	1.31	3	0.2978983	0.7021017	0.393491	-10.718501	1.317202	0.318182
1.37	1.35	4	0.4167635	0.5832365	0.5199559	-10.704733	1.357203	0.439394
1.2	1.37	5	0.4800441	0.5199559	0.5832365	-11.457359	1.395297	0.560606
1.31	1.41	6	0.60650901	0.39349099	0.7021017	-9.3908398	1.435298	0.681818
1.27	1.54	7	0.90509211	0.09490789	0.8025402	-4.1559948	1.482721	0.80303
1.35	1.56	8	0.92938981	0.07061019	0.9209087	-2.3343216	1.555371	0.924242

Technician 2

The p-values are greater than 0.05, both data follows normal distribution

4.14 Two different types of glass bottles are suitable for use by a soft-drink beverage bottler. The internal pressure strength of the bottle is an important quality characteristic. It is known that $\sigma_1 = \sigma_2 = 3.0$ psi. From a random sample of $n_1 = n_2 = 16$ bottles, the mean pressure strengths are observed to be $\bar{x}_1 = 175.8$ psi and $\bar{x}_2 = 181.3$ psi. The company will not use bottle design 2 unless its pressure strength exceeds that of bottle design 1 by at least 5 psi. Based on the sample data, should they use bottle design 2 if we use α=0.05. What is the P-value for this test?

H0 : $\mu_2 - \mu_1 \leq 5$ H1 : $\mu_2 - \mu_1 > 5$

$\alpha = 0.05$

$((\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1) / \sigma \sqrt{(1/n_1 + 1/n_2)})$

$= ((181.3 - 175.8) - 5)/3\sqrt{(1/16 + 1/16)}$

$=0.471$

p-value = 0.3189 > 0.05

fail to reject H0

4.19 An inspector counts the surface-finish defects in dishwashers. A random sample of five dishwashers contains three such defects. Is there reason to conclude that the mean occurrence rate of surface-finish defects per dishwasher exceeds 0.5? Use the results of part (a) of Exercise 4.18 and assume that α=0.05.

H0 : $p \leq 0.5$ H1 : $p > 0.5$

$\hat{p} = 3/5 = 0.6$ $n = 5$ $P_0 = 0.5$

$Z_{0.05} = 1.644$

$z = (\hat{P} - P_0)/\sqrt{(P_0*(1-P_0)/n)} = (0.6-0.5)/(0.5*0.5/5) = 2 > 1.644$

reject H0, we can say that the mean occurrence rate of surface-finish defects per dishwasher exceeds 0.5

4.24 Plot the residuals from Exercise 4.23 against the firing temperatures. Is there any indication that variability in baked anode density depends on the firing temperature? What firing temperature would you recommend using?

單因子變異數分析

摘要

組	個數	總和	平均	變異數
500	6	250.2	41.7	0.02
525	6	249.5	41.58333	0.037667
550	6	248.7	41.45	0.115
575	6	248	41.33333	0.246667

ANOVA

變源	SS	自由度	MS	F	P-值	臨界值
組間	0.456667	3	0.152222	1.45204	0.257611	3.098391
組內	2.096667	20	0.104833			
總和	2.553333	23				

p-value = 0.258 > 0.05

fail to reject H0, we cannot say that variability in baked anode density depends on the firing temperature

I recommend to use 500 fire temperature, because its variance is the smallest, which means it's steady.

4.27 A plant distills liquid air to produce oxygen, nitrogen, and argon. The percentage of impurity in the oxygen is thought to be linearly related to the amount of impurities in the air as measured by the "pollution count" in parts per million (ppm). A sample of plant operating data is shown below:

Purity (%)		93.3	92.0	92.4	91.7	94.0	94.6	93.6
Pollution count (ppm)		1.10	1.45	1.36	1.59	1.08	0.75	1.20
Purity (%)	93.1	93.2	92.9	92.2	91.3	90.1	91.6	91.9
Pollution count (ppm)	0.99	0.83	1.22	1.47	1.81	2.03	1.75	1.68

摘要輸出

迴歸統計	
R 的倍數	0.934851
R 平方	0.873946
調整的 R -	0.86425
標準誤	0.137843
觀察值個	15

ANOVA

	自由度	SS	MS	F	顯著值
迴歸	1	1.71255	1.71255	90.13057	3.28E-07
殘差	13	0.24701	0.019001		
總和	14	1.95956			

	係數	標準誤	t 統計	P-值	下限 95%	上限 95%	下限 95.0%	上限 95.0%
截距	29.2287	2.936338	9.954133	1.9E-07	22.88513	35.57228	22.88513	35.57228
X 變數 1	-0.30126	0.031733	-9.49371	3.28E-07	-0.36982	-0.23271	-0.36982	-0.23271

(a) Fit a linear regression model to the data.

$Y = \beta_0 + \beta_1 x = 29.23 - 0.3x$

(b) Test for significance of regression.

H0 : $\beta_1 = 0$ H1 : $\beta_1 \neq 0$

t-test = 90.13

p-value = 0.0001 < 0.05

reject H0, we can say that there is a significant difference between purity and pollution count.

(c) Find a 95% confidence interval on β_1.

根據上表

[-0.37, -0.23]