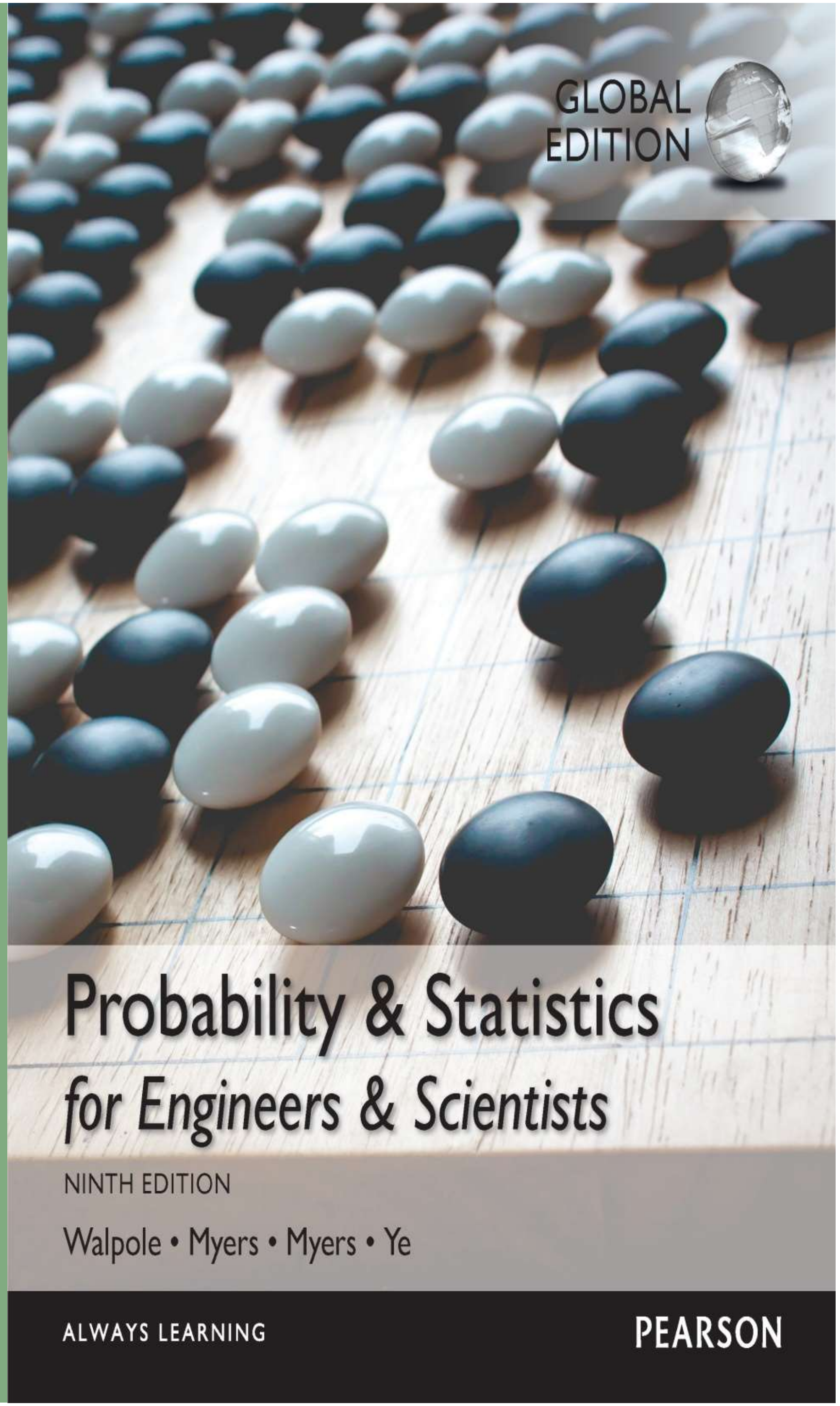
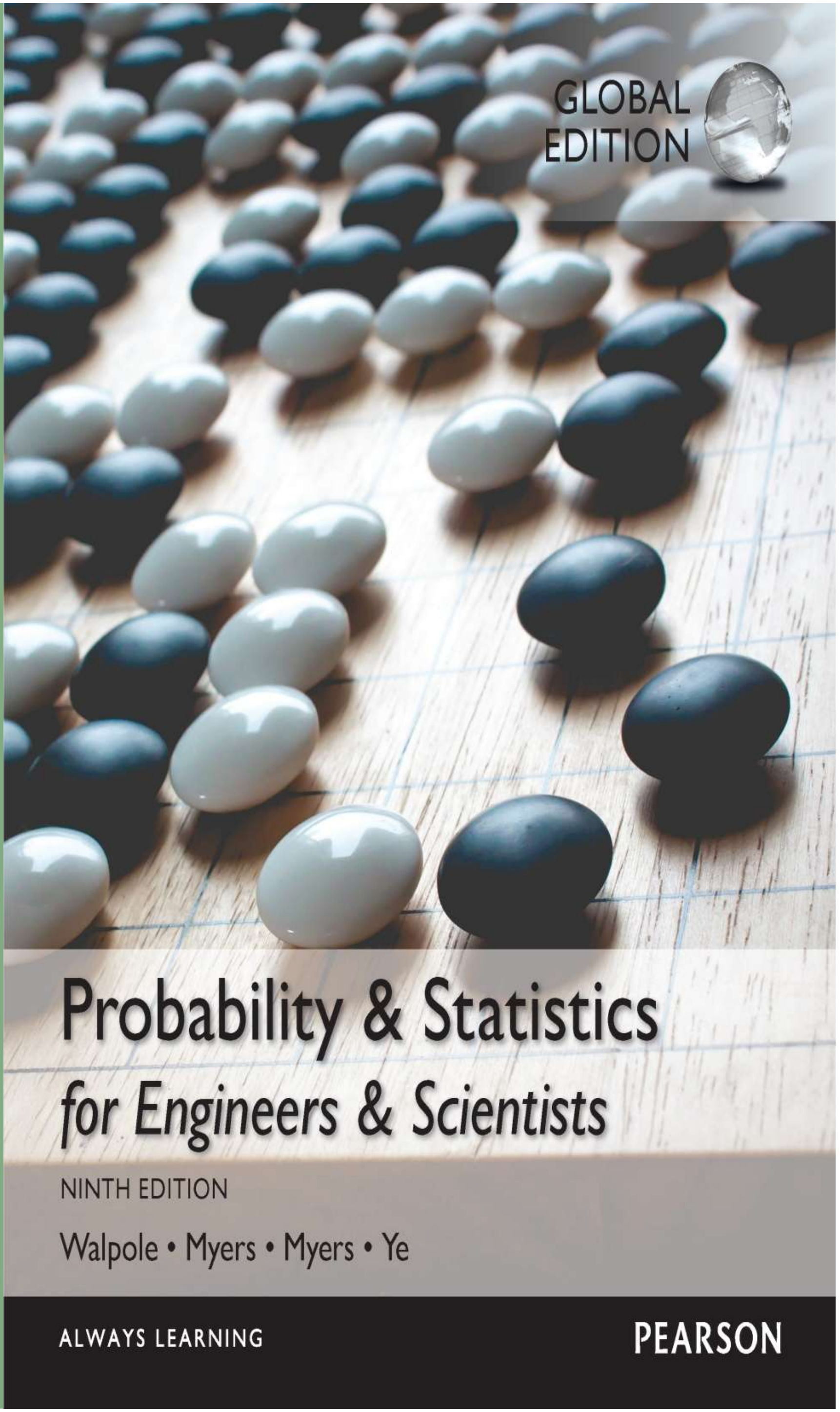
Chapter 8

Fundamental Sampling Distributions and Data Descriptions



Section 8.7

F-Distribution



Theorem 8.6

Let U and V be two independent random variables having chi-squared distributions with v_1 and v_2 degrees of freedom, respectively. Then the distribution of the random variable $F = \frac{U/v_1}{V/v_2}$ is given by the density function

$$h(f) = \begin{cases} \frac{\Gamma[(v_1+v_2)/2](v_1/v_2)^{v_1/2}}{\Gamma(v_1/2)\Gamma(v_2/2)} \frac{f^{(v_1/2)-1}}{(1+v_1f/v_2)^{(v_1+v_2)/2}}, & f > 0, \\ 0, & f \leq 0. \end{cases}$$

This is known as the ***F*-distribution** with v_1 and v_2 degrees of freedom (d.f.).

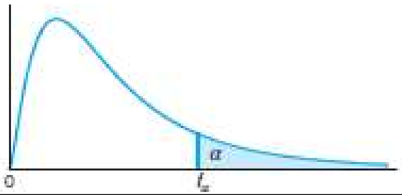


Table A.6 Critical Values of the F-Distribution

	$f_{0.05}(v_1, v_2)$								
	v_1								
v_2	1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65

Theorem 8.7

Writing $f_\alpha(v_1, v_2)$ for f_α with v_1 and v_2 degrees of freedom, we obtain

$$f_{1-\alpha}(v_1, v_2) = \frac{1}{f_\alpha(v_2, v_1)}.$$

Theorem 8.8

If S_1^2 and S_2^2 are the variances of independent random samples of size n_1 and n_2 taken from normal populations with variances σ_1^2 and σ_2^2, respectively, then

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2}$$

has an F-distribution with $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$ degrees of freedom.

What Is the *F*-Distribution Used For?

Paint	Sample Mean	Sample Variance	Sample Size
A	$\bar{X}_A = 4.5$	$s_A^2 = 0.20$	10
B	$\bar{X}_B = 5.5$	$s_B^2 = 0.14$	10
C	$\bar{X}_C = 6.5$	$s_C^2 = 0.11$	10

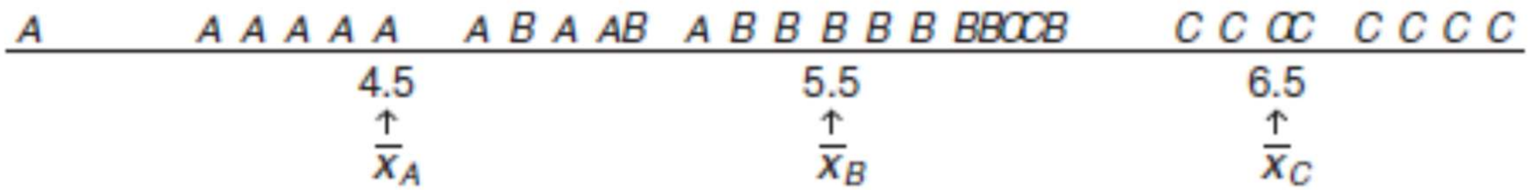


Figure 8.13: Data from three distinct samples.

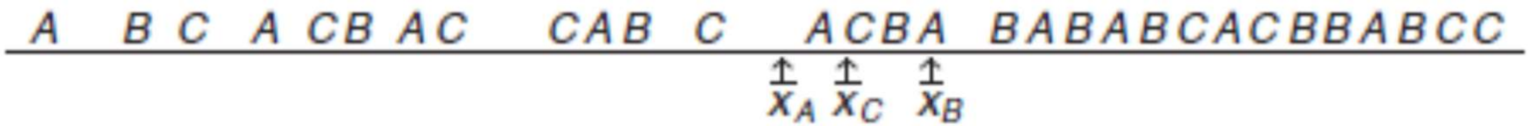


Figure 8.14: Data that easily could have come from the same population.

Related examples

1.A soft-drink machine is regulated so that the amount of drink dispensed averages 240 milliliters with a standard deviation of 15 milliliters. Periodically, the machine is checked by taking a sample of 40 drinks and computing the average content. If the mean of the 40 drinks is a value within the interval $\mu_{\bar{x}} \pm 2\sigma_{\bar{x}}$, the machine is thought to be operating satisfactorily; otherwise, adjustments are made. the company official found the mean of 40 drinks to be $\bar{x}$= 236 milliliters and concluded that the machine needed no adjustment. Was this a reasonable decision?

2-The distribution of the weights of one-week-old chicks of a certain breed has a mean of 120 grams and a standard deviation of 12 grams. The distribution of weights of chicks of the same age of another breed is 92 grams with a standard deviation of 8 grams. Assuming that the sample mean can be calculated to any degree of accuracy, find the probability that the sample mean for the weights of 46 random chicks of first breed exceeds the sample mean for the weights of 84 random chicks of the second breed by at most 20 grams.

3-For a chi-square distribution, find χ^2_α such that

(a) $P(\chi^2 > \chi^2_\alpha) = 0.95$, when $v = 5$;

(b) $P(\chi^2 > \chi^2_\alpha) = 0.99$, when $v = 13$;

(c) $P(19.68 < \chi^2 < \chi^2_\alpha) = 0.04$, when $v = 11$.

4-The scores on a placement test given to college freshmen for the past five years are approximately normally distributed with a mean $\mu = 74$ and a variance $\sigma^2 = 8$. Would you still consider $\sigma^2 = 8$ to be a valid value of the variance if a random sample of 20 students who take the placement test this year obtain a value of $s^2 = 20$? (α=0.01)

5-A transport company claims that the average running time for a bus on a particular route is 300 minutes. Six buses are randomly observed and their running times are recorded as follows: 320, 310, 295, 312, 302, and 308 minutes. Would you agree with the transport company’s claim? Assume a normal distribution.

6-Pull-strength tests on 10 soldered leads for semiconductor device yield the following results, in pounds of force required to rupture the bond:

19.8 12.7 13.2 16.9 10.6 18.8 11.1 14.3 17.0 12.5

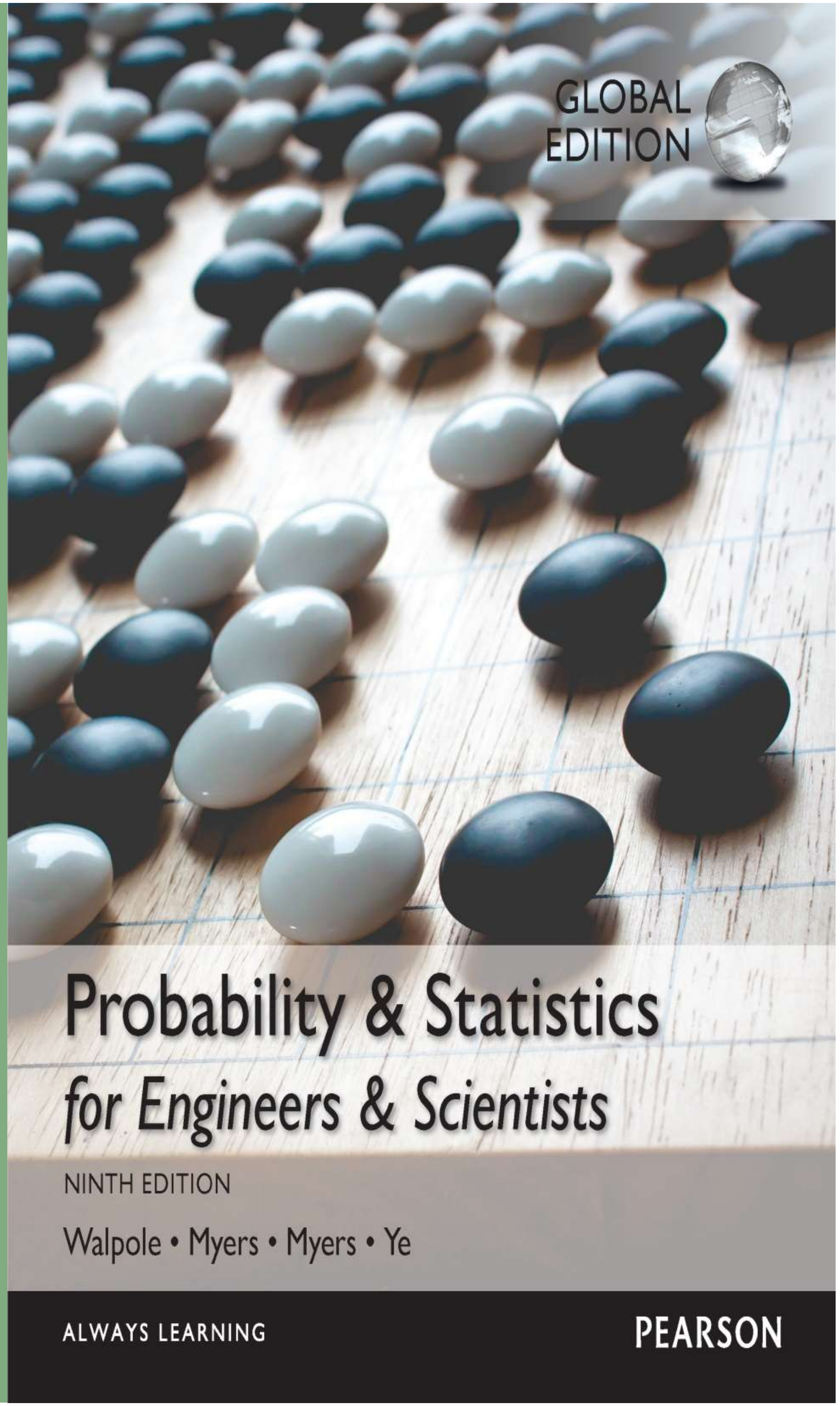
Another set of 8 leads was tested after encapsulation to determine whether the pull strength had been increased by encapsulation of the device, with the following results:

24.9 22.8 23.6 22.1 20.4 21.6 21.8 22.5

Comment on the evidence available concerning equality of the two population variances.

Section 8.8

Quantile and Probability Plots



Definition 8.6

A **quantile** of a sample, $q(f)$, is a value for which a specified fraction f of the data values is less than or equal to $q(f)$.

A **quantile plot** simply *plots the data values on the vertical axis against an empirical assessment of the fraction of observations exceeded by the data value.* For theoretical purposes, this fraction is computed as

$$f_i = \frac{i - \frac{3}{8}}{n + \frac{1}{4}},$$

where i is the order of the observations when they are ranked from low to high. In other words, if we denote the ranked observations as

$$y_{(1)} \leq y_{(2)} \leq y_{(3)} \leq \cdots \leq y_{(n-1)} \leq y_{(n)},$$

then the quantile plot depicts a plot of $y_{(i)}$ against f_i.

Table 1.9: Data for Example 1.6

Sample	Measurements	Sample	Measurements
1	29 36 39 34 34	16	35 30 35 29 37
2	29 29 28 32 31	17	40 31 38 35 31
3	34 34 39 38 37	18	35 36 30 33 32
4	35 37 33 38 41	19	35 34 35 30 36
5	30 29 31 38 29	20	35 35 31 38 36
6	34 31 37 39 36	21	32 36 36 32 36
7	30 35 33 40 36	22	36 37 32 34 34
8	28 28 31 34 30	23	29 34 33 37 35
9	32 36 38 38 35	24	36 36 35 37 37
10	35 30 37 35 31	25	36 30 35 33 31
11	35 30 35 38 35	26	35 30 29 38 35
12	38 34 35 35 31	27	35 36 30 34 36
13	34 35 33 30 34	28	35 30 36 29 35
14	40 35 34 33 35	29	38 36 35 31 31
15	34 35 38 35 30	30	30 34 40 28 30

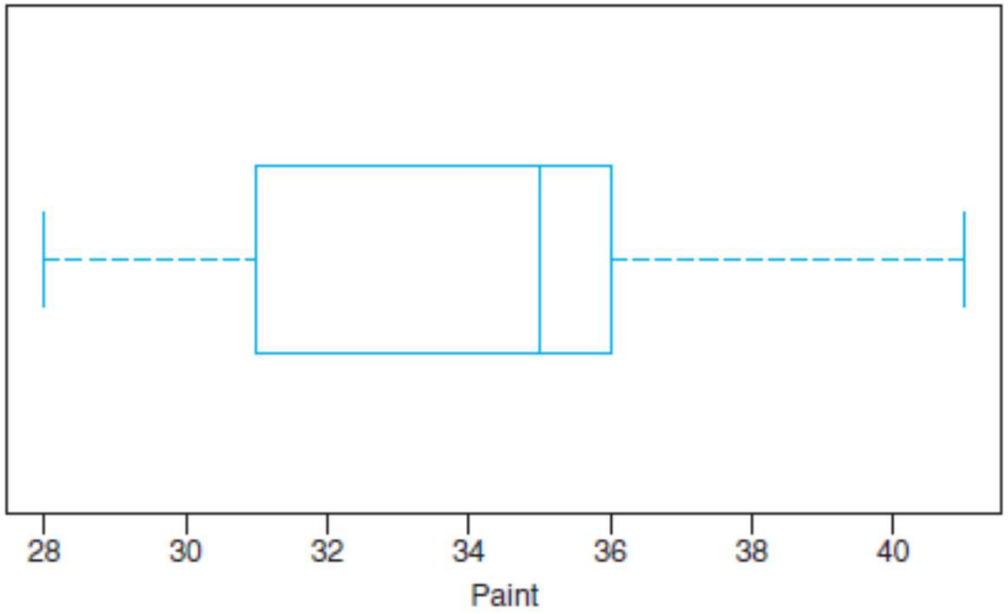
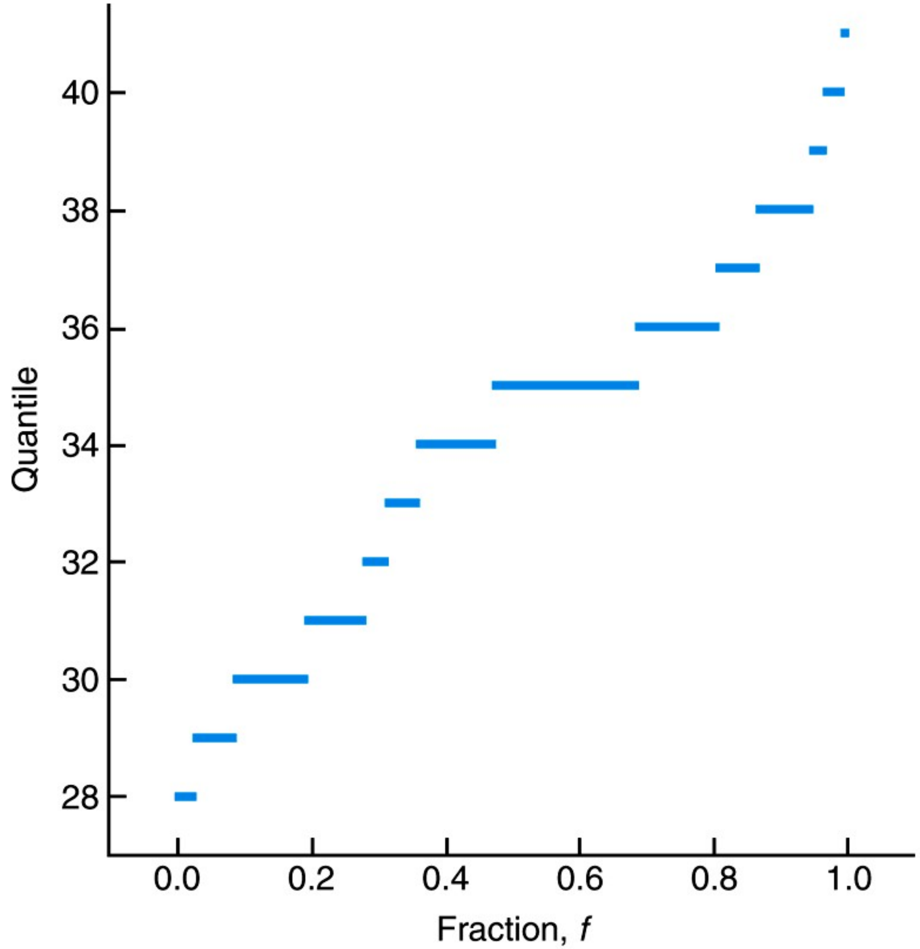


Figure 1.11: Box-and-whisker plot for thickness of paint can "ears."

Figure 8.15 Quantile plot for paint data



Definition 8.7

The **normal quantile-quantile plot** is a plot of $y_{(i)}$ (ordered observations) against $q_{0,1}(f_i)$, where $f_i = \frac{i-\frac{3}{8}}{n+\frac{1}{4}}$.

$$q_{\mu,\sigma}(f) = \mu + \sigma\{4.91[f^{0.14} - (1-f)^{0.14}]\}.$$

$$q_{0,1}(f) = 4.91[f^{0.14} - (1-f)^{0.14}].$$

Figure 8.16 Normal quantile-quantile plot for paint data

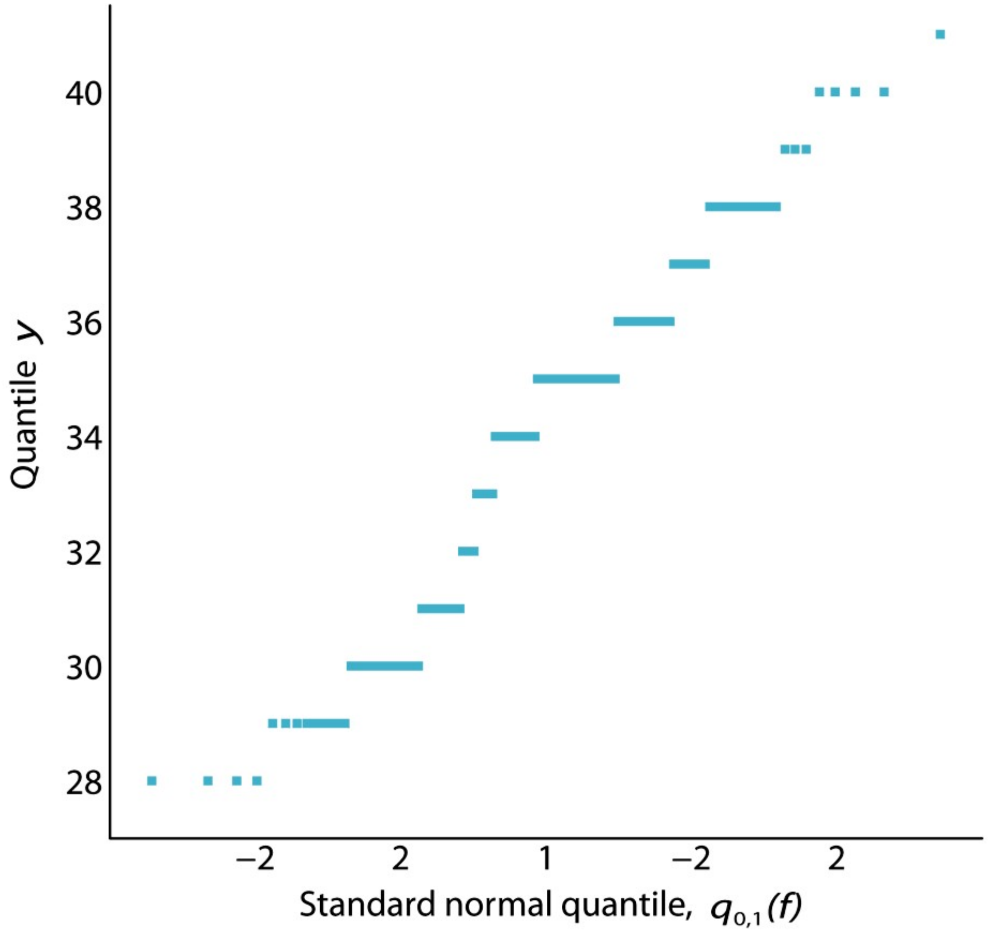


Table 8.1 Data for Example 8.12

Number of Organisms per Square Meter			
Station 1		Station 2	
5,030	4,980	2,800	2,810
13,700	11,910	4,670	1,330
10,730	8,130	6,890	3,320
11,400	26,850	7,720	1,230
860	17,660	7,030	2,130
2,200	22,800	7,330	2,190
4,250	1,130		
15,040	1,690		

Figure 8.17 Normal quantile-quantile plot for density data of Example 8.12