Test Your Understanding 0

The following table shows counts of vehicles passing a particular point on a road, grouped by day (columns) and hour (rows). There is something wrong with the table. Can you tell what it is?

	Day						
Hours	Mon	Tue	Wed	Thu	Fri	Sat	Sun
0000-0100	41	10	12	5	15	19	35
0100-0200	29	4	5	5	3	8	10
0200-0300	17	3	3	0	4	6	9
0300-0400	7	5	5	3	5	3	4
0400-0500	10	24	20	20	20	26	18
0500-0600	20	60	63	62	70	68	50
0600-0700	42	219	232	220	225	222	89
0700-0800	91	429	425	473	447	446	207
0800-0900	163	270	274	272	285	281	314
0900-1000	222	201	206	190	210	227	391
1000-1100	276	224	227	233	235	236	395
1100-1200	280	261	240	262	281	272	398
1200-1300	321	350	337	309	321	372	389
1300-1400	193	234	231	251	232	293	404
1400-1500	202	240	299	282	317	313	353
1500-1600	226	315	316	299	322	402	320
1600-1700	203	286	309	293	339	381	301
1700-1800	220	405	377	393	386	362	281
1800-1900	206	303	295	270	302	299	213
1900-2000	193	257	238	242	254	242	219
2000-2100	123	128	170	153	177	156	122
2100-2200	99	106	111	109	102	98	79
2200-2300	74	66	95	85	82	86	66
2300-2400	54	39	54	50	59	86	80

Test Your Understanding 1

A process that produces audiotapes is monitored for the thickness of magnetic coating on the tapes. In Figure 1, the thicknesses of 150 thickness measurements (in microns) are plotted versus the order in which they were taken. Your boss asks you how the process is performing. What do you say?
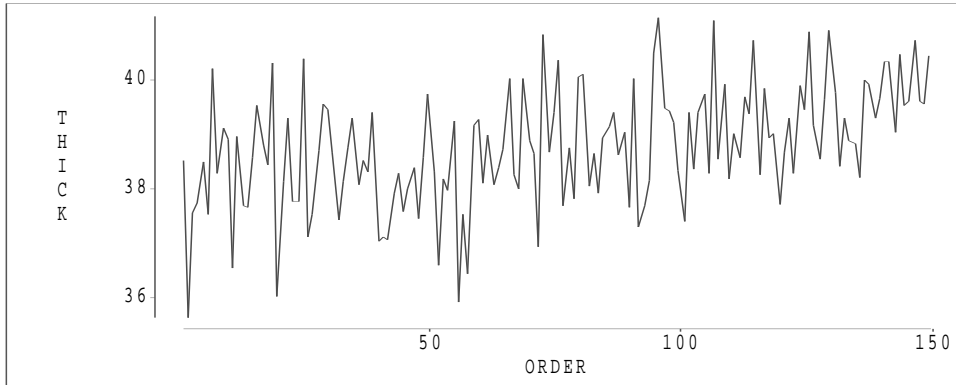

Figure 1: *One-hundred fifty consecutive measurements of the thickness of magnetic coating on audio tape*

Figure 2 shows a time series plot. What is the simplest moving average that will remove the cycles in the plot? Apply the moving average to the first 5 data values: -0.95, 0.79, -0.99, 1.20, -1.03. Plot the moving average values on the graph to demonstrate the cycles have been removed.
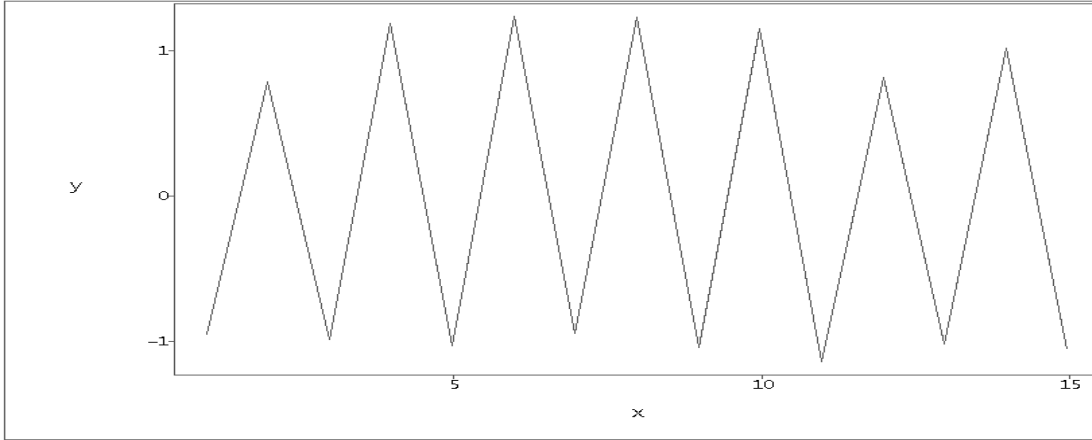

Figure 2: *Time series plot*

Test Your Understanding 3

A microhardness tester is a machine designed to test the hardness of material. In order to test the repeatability and reproducibility of the measuring process using a particular microhardness tester, four operators each took 20 measurements of the hardness of the same metal piece. Time series plots showed that the measuring process was stationary for each of them. Summarize what the stratified plot in Figure 3 tells about the R&R of the measuring process.
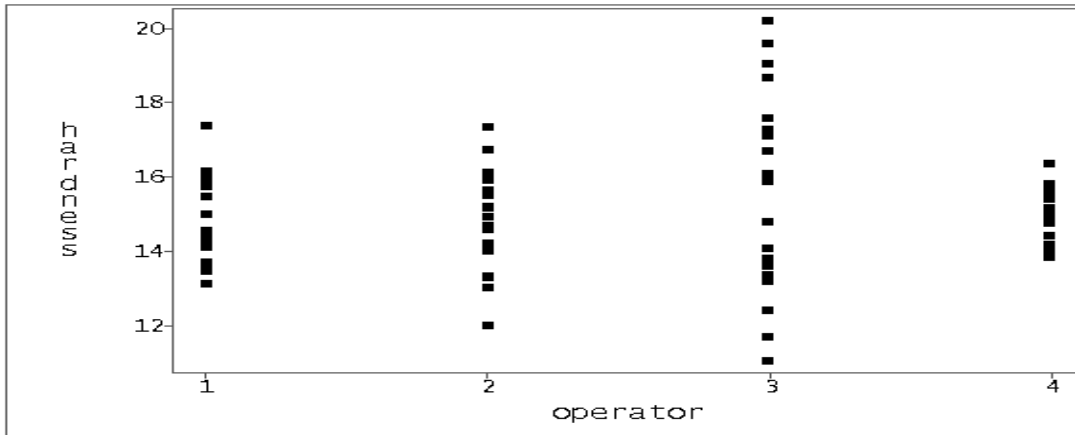

Figure 3: *Force measurements from a microhardness tester*

Figure 3 shows four frequency histograms.
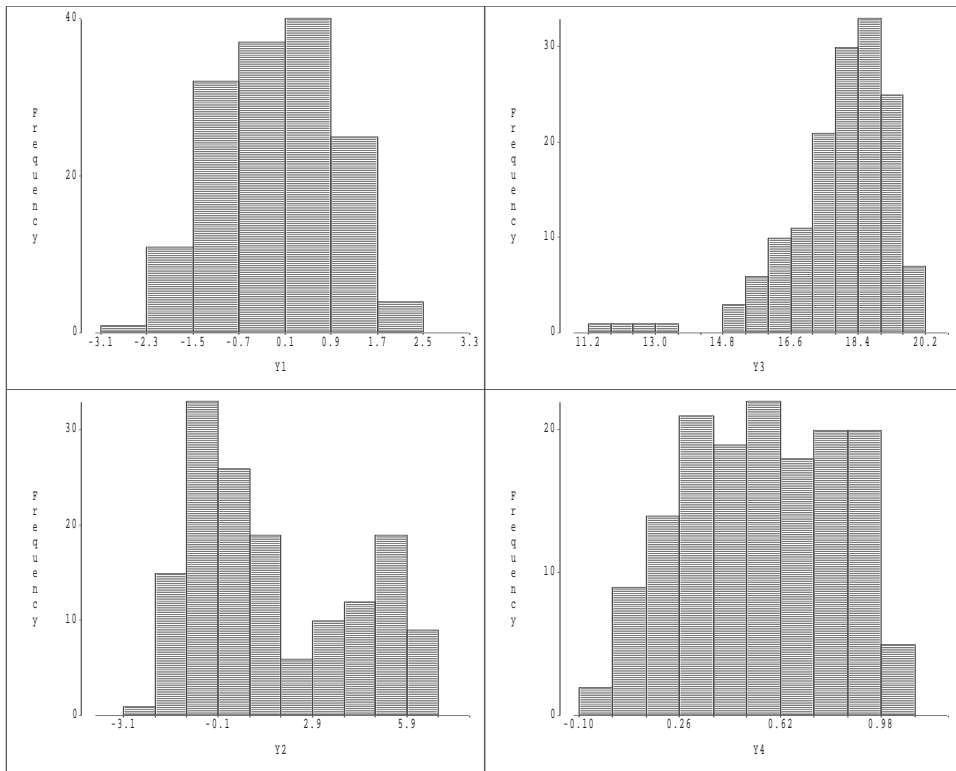

Figure 4: *Four frequency histograms*

Briefly describe the main features of each histogram.
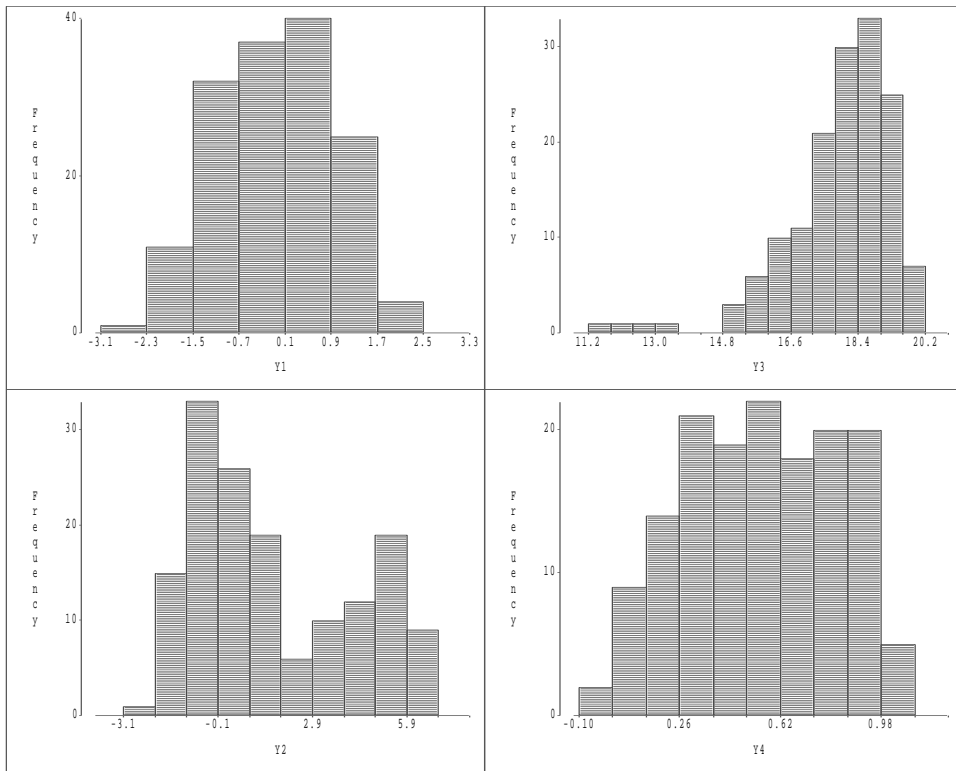
Figure 3 shows four frequency histograms.

Figure 5: *Four frequency histograms*

For each histogram, describe the summary measures you would use to back up your description.

The lengths of seven telephone calls, in minutes, are 17, 7, 1, 4, **39**, 2, 11. Generate a boxplot for these data. Does the boxplot identify any outliers?

Test Your Understanding 7

Compute a 1-time trimmed mean and a 1-time Winsorized mean for the data from the last TYU: 17, 7, 1, 4, 39, 2, 11

Test Your Understanding 8

Suppose you want to estimate the number of figures in a 10 chapter textbook, but that you can only sample 20 pages.

1. If you believe all chapters have roughly the same distribution of figures, how would you choose the 20 pages?

2. If you believe that chapters 3, 4, 7 and 9 have many more figures per page than the others, how might you choose the 20 pages?

Test Your Understanding 9

To compare the efficacy of mosquito repellant, volunteers have an arm coated with a prescribed amount of the product. The arm is then inserted into a chamber filled with mosquitoes for a fixed amount of time and the number of bites counted (YUCK!). To compare the efficacies of two different repellants, volunteers are randomly divided into two groups. One group is given repellant 1 and the other repellant 2 and the test described above is conducted for each.

1. Is this a controlled experiment? Why?

2. If it is a controlled experiment, describe the

   (a) Experimental units

   (b) Response

   (c) Experimental factor(s)

   (d) Possible nuisance factors

   (e) Factor levels

   (f) Treatments

   (g) Effect

Recall the experiment described in TYU 9:

To compare the efficacy of mosquito repellant, volunteers have an arm coated with a prescribed amount of the product. The arm is then inserted into a chamber filled with mosquitoes for a fixed amount of time and the number of bites counted (YUCK!). To compare the efficacies of two different repellants, volunteers are randomly divided into two groups. One group is given repellant 1 and the other repellant 2 and the test described above is conducted for each.

How could blocking be used to improve the design?

Test Your Understanding 11

In order to identify risk factors for juvenile crimimal behavior, researchers compared a large group of juvenile offenders with a group of their peers who were not offenders. These groups were compared with respect to a large number of factors.

1. What kind of a study is this? Be as specific as you can.

2. Suppose the researchers find a number of factors that are quite different for the two groups. Do you think the researchers can conclude these factors cause juvenile crime? Explain.

Test Your Understanding 12

Experience has shown that the width, in mm, of the flange on a plastic connector has the following distribution:

$$p_Y(y) = 50y, \ 0.48 < y < 0.52,$$
$$= 0, \ \text{otherwise}$$

(a) Of the next 1000 connectors produced, how many do you estimate will have widths between 0.50 and 0.51 mm? Show how you arrived at your estimate.

(b) How many times as likely is it to produce connectors with flange width close to 0.51 mm as it is to produce connectors with flange width close to 0.49 mm? Justify your answer.

Test Your Understanding 13

Suppose we sample 4 OJ containers from the production lot having population proportion p of acceptable containers. Calculate $p_Y(3) = P(Y = 3)$, the probability of obtaining exactly 3 acceptable containers in the sample.

Test Your Understanding 14

A system consists of three identical components. The system can operate successfully only if at least two components are operating. The probability any one component lasts less than 100 hours is 0.06, and whether that component fails before 100 hours is independent of the performance of the other two components. If Y is the number of components in the system that fail before 100 hours,

(a) Obtain the distribution model of Y.

(b) What is the probability the system fails before 100 hours?

Test Your Understanding 15

Suppose the population of math SAT scores follows a normal distribution with mean 500 and standard deviation 80. What proportion of students get between 600 and 700 on the exam?

Test Your Understanding 16

The state bar exam is designed so that 30% of prospective lawyers pass it each year, and over time, this passing percentage has held true, on average. From one year to the next, however, the percentage can vary. If 1000 prospective lawyers take the exam this year, and assuming each lawyer has a 0.30 probability of passing, approximately what is the probability that 320 or more pass?

Test Your Understanding 17

Fill in the remainder of the table and create the normal quantile plot. Remember the data are 266, 149, 161, 220.

k	$(k-0.375)/(n+0.250)$	$q_{(k)}$
1	0.147	-1.05
2	0.382	-0.30
3	.	.
4	.	.

Test Your Understanding 18

A data set consists of values 10, 12 and 23. It is believed that the C+E model is appropriate for these data, and it is desired to estimate the population mean, μ, based on these data.

a. In terms of SSE, which of the two estimates, $\hat{\mu}_1 = 12$ or $\hat{\mu}_2 = 14$ is better?

b. What is the least squares estimate of μ? Show that it has a better SSE value than the other two estimates.

Test Your Understanding 19

20 tires of a new design give tread lives having a sample mean of 42,357 miles and a sample standard deviation of 3,042 miles. The data show no evidence of outliers or non-normality.

a. Obtain a 95% confidence interval for the true mean tread life.

b. From this interval can you conclude that mean tread life exceeds 42,000 miles? Why or why not?

Test Your Understanding 20

Recall the problem from Test Your Understanding 19: *Tests of 20 tires of a new design give tread lives having a sample mean of 42,357 miles and a sample standard deviation of 3,042 miles. The data show no evidence of outliers or non-normality.*

a. Obtain a 95% prediction interval interval for a new observation.

b. A customer complains that he bought a set of four tires of this type and got only 38,000 miles on one of them. Based on your interval from part a, do you think this is a defective tire? Why?

Ecologists want to estimate the proportion of deer in a certain region infected with the Lyme disease bacterium, Borrelia burgdorfen. To do so, they collect a random sample of 25 deer. Analysis shows that 6 of the deer are carrying the bacterium. Assuming the number of deer in the region is large, compute the approximate score level 0.90 confidence interval for the proportion of infected deer in the region.

Recall the problem from Test Your Understanding 19: *Tests of 20 tires of a new design give tread lives having a sample mean of 42,357 miles and a sample standard deviation of 3,042 miles. The data show no evidence of outliers or non-normality.*

Suppose we want to use a level 0.95 interval to estimate mean tread life to within 500 miles. We are willing to assume that $\sigma = 3000$. How large a sample will be needed?

Test Your Understanding 23

In a study to determine if there is a difference in gripping strength between the dominant and nondominant hand, 71 students measured the strengths of their grips with each hand. The mean of the differences between the dominant and nondominant hand measurements was 3.73 and the standard deviation was 4.83 pounds. Compute a 99% confidence interval to decide if there is a difference in mean grip between the dominant and nondominant hand. What do you conclude?

Test Your Understanding 24

Recall the data from Test Your Understanding 19: *Tests of 20 tires of a new design give tread lives having a sample mean of 42,357 miles and a sample standard deviation of 3,042 miles. The data show no evidence of outliers or non-normality.* A competitor introduces a new tire. In independent tests, 12 of the competitor's tires gave tread lives having a sample mean of 44,821 miles and a sample standard deviation of 2,568 miles. These data also show no evidence of outliers or non-normality. Assuming the population standard deviations are equal,

a. Obtain a 90% confidence interval for the difference of mean tire life for the two tire types.

Solution: $L = 0.90$, $\overline{y}_1 = 42357$, $s_1 = 3042$, $n_1 = 20$, $\overline{y}_2 = 44821$, $s_2 = 2568$, $n_2 = 12$,

Therefore,
$$t_{n_1+n_2-2,(1+L)/2} = t_{30,0.95} = 1.6973,$$

and
$$s_p^2 = \frac{(20-1)3042^2 + (12-1)2568^2}{20+12-2} = 8278740.2$$

The estimated standard error of $\overline{Y}_1 - \overline{Y}_2$ *is*

$$\hat{\sigma}(\overline{Y}_1 - \overline{Y}_2) = \sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{8278740.2\left(\frac{1}{20} + \frac{1}{12}\right)} = 1050.63$$

The interval is
$$42357 - 44821 \pm (1.6973)(1050.63) = (-4247, -681)$$

b. Based on this interval, can you conclude the mean tire lives for the two tire types are different? Why?

Solution: *Yes, we can conclude the mean tire lives for the two tire types are different, since the interval does not contain 0. In fact, with 90% confidence, we estimate that the mean tire life for tire 2 is between 681 and 4247 miles greater than that for tire 1.*