SECTION 7.3 – Assessing Normality

Normal Probability Plot

A normal probability plot is a graph that plots observed data versus normal scores. A normal score is the expected z-score of the data value, assuming that the distribution of the random variable is normal. The expected z-score of an observed value depends on the number of observations in the data set.

Drawing a Normal Probability Plot

Step 1 – Arrange the data in ascending order.

Step 2 – Compute $f_i = \dfrac{i - 0.375}{n + 0.25}$, where i is the index (the position of the data value in the ordered list) and n is the number of observations. The expected proportion of observations less than or equal to the ith data value is f_i.

Step 3 – Find the z-score corresponding to f_i from Table V.

Step 4 – Plot the observed values on the horizontal axis and the corresponding expected z-scores on the vertical axis.

If sample data are taken from a population that is normally distributed, a normal probability plot of the observed values versus the expected z-scores will be approximately linear.

Conclusions based off a normal probability plot

To assess the normality of a variable using sample data, construct a normal probability plot.

- If the plot is roughly linear, you can assume that the variable is approximately normally distributed.
- If the plot is not roughly linear, you can assume that the variable is not approximately normally distributed.

These guidelines should be interpreted loosely for small samples but usually interpreted strictly for large samples.

Note: Your text will have a different way of proving whether the plot is linear enough rather than by appearance, especially if it is not obvious one way or another. The author takes it a step further by having you calculate the correlation coefficient, r, from Section 4.1 and then comparing it with a critical value from a new table, Table VI. So, to keep things simpler, we will just use the guidelines from above.

SECTION 7.3 – Assessing Normality

☺ **Exercises:**

1) A simple random sample of 15 resting heart rates of students from a journalism class is shown below:

92, 73, 61, 80, 72, 75, 79, 83, 82, 63, 95, 71, 65, 89, 74

Construct a normal probability plot for the data and use the plot to assess the normality of the students' heart rates.

Solution →

To construct a normal probability plot, we create a 4-column table as shown below. Since there are 15 observations, $n = 15$.

Column 1: Index, i. Write the numbers 1, 2, 3, …, n in column 1, where n is the number of observations)

Column 2: We arrange the data in increasing order and place in column 2. These are your x-values for the plot.

Column 3: Compute $f_i = \dfrac{i - 0.375}{n + 0.25}$, where i is the index in column 1 and n is the number of observations.

Column 4: Find the z-score corresponding to f_i in column 3 from Table V. These are your y-values for the plot.

	x		***y***
Index, i	**Observed Values**	f_i	**Expected z-score**

SECTION 7.3 – Assessing Normality

☺ **Exercises:** Here is an alternate faster method to the exercise on the previous page.

1) A simple random sample of 15 resting heart rates of students from a journalism class is shown below:

92, 73, 61, 80, 72, 75, 79, 83, 82, 63, 95, 71, 65, 89, 74

Construct a normal probability plot for the data and use the plot to assess the normality of the students' heart rates.

Solution →

The method on the previous page consists of creating a 4-column table in which one must calculate the n expected z-scores from each corresponding f_i value. So, that means for any data set with a sample size of 15, will create the exact same collection of z-scores. So, we can create a table of these normal scores for sample sizes ranging from 5 to 30 inclusive and bypass all of the calculations. A custom "Table of Normal Scores" is located on the next page.
(*Note*: Table VI in your text contains the critical values. The table on the following page has been created to contain the normal scores instead of the student generating them from Table V. We will introduce this table as The Table of Normal Scores. This table is not in your textbook.)

So, now, to construct a normal probability plot, we will just create a 2-column table as shown below.

Column 1: We arrange the data in increasing order and place in column 1. These are your x-values for the plot.

Column 2: Since there are 15 observations, we will use the normal scores from the $n = 15$ column from the Table of Normal Scores on the next page and place those values in column 2.

x	y
Observed Values	**Normal Scores**

Since the normal probability plot is roughly linear, then one can conclude that the resting heart rates for students is approximately normally distributed.

SECTION 7.3 – Assessing Normality

Table of Normal Scores (Custom)

Ordered Position	*n*												
	5	***6***	***7***	***8***	***9***	***10***	***11***	***12***	***13***	***14***	***15***	***16***	***17***
1	–1.18	–1.28	–1.36	–1.43	–1.50	–1.55	–1.59	–1.64	–1.68	–1.71	–1.74	–1.77	–1.80
2	–0.50	–0.64	–0.76	–0.85	–0.93	–1.00	–1.06	–1.11	–1.16	–1.20	–1.24	–1.28	–1.32
3	0.00	–0.20	–0.35	–0.47	–0.57	–0.65	–0.73	–0.79	–0.85	–0.90	–0.94	–0.99	–1.03
4	0.50	0.20	0.00	–0.15	–0.27	–0.37	–0.46	–0.53	–0.60	–0.66	–0.71	–0.76	–0.80
5	1.18	0.64	0.35	0.15	0.00	–0.12	–0.22	–0.31	–0.39	–0.45	–0.51	–0.57	–0.62
6		1.28	0.76	0.47	0.27	0.12	0.00	–0.10	–0.19	–0.27	–0.33	–0.39	–0.45
7			1.36	0.85	0.57	0.37	0.22	0.10	0.00	–0.09	–0.16	–0.23	–0.29
8				1.43	0.93	0.65	0.46	0.31	0.19	0.09	0.00	–0.08	–0.15
9					1.50	1.00	0.73	0.53	0.39	0.27	0.16	0.08	0.00
10						1.55	1.06	0.79	0.60	0.45	0.33	0.23	0.15
11							1.59	1.11	0.85	0.66	0.51	0.39	0.29
12								1.64	1.16	0.90	0.71	0.57	0.45
13									1.68	1.20	0.94	0.76	0.62
14										1.71	1.24	0.99	0.80
15											1.74	1.28	1.03
16												1.77	1.32
17													1.80

Ordered Position	*n*												
	18	***19***	***20***	***21***	***22***	***23***	***24***	***25***	***26***	***27***	***28***	***29***	***30***
1	–1.82	–1.85	–1.87	–1.89	–1.91	–1.93	–1.95	–1.97	–1.98	–2.00	–2.01	–2.03	–2.04
2	–1.35	–1.38	–1.40	–1.43	–1.45	–1.48	–1.50	–1.52	–1.54	–1.56	–1.58	–1.59	–1.61
3	–1.06	–1.10	–1.13	–1.16	–1.18	–1.21	–1.24	–1.26	–1.28	–1.30	–1.32	–1.34	–1.36
4	–0.84	–0.88	–0.92	–0.95	–0.98	–1.01	–1.04	–1.06	–1.09	–1.11	–1.13	–1.15	–1.17
5	–0.66	–0.70	–0.74	–0.78	–0.81	–0.84	–0.87	–0.90	–0.93	–0.95	–0.98	–1.00	–1.02
6	–0.50	–0.54	–0.59	–0.63	–0.66	–0.70	–0.73	–0.76	–0.79	–0.82	–0.84	–0.87	–0.89
7	–0.35	–0.40	–0.45	–0.49	–0.53	–0.57	–0.60	–0.63	–0.66	–0.69	–0.72	–0.75	–0.77
8	–0.21	–0.26	–0.31	–0.36	–0.40	–0.44	–0.48	–0.52	–0.55	–0.58	–0.61	–0.64	–0.67
9	–0.07	–0.13	–0.19	–0.24	–0.28	–0.33	–0.37	–0.41	–0.44	–0.48	–0.51	–0.54	–0.57
10	0.07	0.00	–0.06	–0.12	–0.17	–0.22	–0.26	–0.30	–0.34	–0.38	–0.41	–0.44	–0.47
11	0.21	0.13	0.06	0.00	0.06	–0.11	–0.15	–0.20	–0.24	–0.28	–0.31	–0.35	–0.38
12	0.35	0.26	0.19	0.12	0.06	0.00	–0.05	–0.10	–0.14	–0.18	–0.22	–0.26	–0.29
13	0.50	0.40	0.31	0.24	0.17	0.11	0.05	0.00	–0.05	–0.09	–0.13	–0.17	–0.21
14	0.66	0.54	0.45	0.36	0.28	0.22	0.15	0.10	0.05	0.00	–0.04	–0.09	–0.12
15	0.84	0.70	0.59	0.49	0.40	0.33	0.26	0.20	0.14	0.09	0.04	0.00	–0.04
16	1.06	0.88	0.74	1.63	0.53	0.44	0.37	0.30	0.24	0.18	0.13	0.09	0.04
17	1.35	1.10	0.92	1.78	0.66	0.57	0.48	0.41	0.34	0.28	0.22	0.17	0.12
18	1.82	1.38	1.13	0.95	0.81	0.70	0.60	0.52	0.44	0.38	0.31	0.26	0.21
19		1.85	1.40	1.16	0.98	0.84	0.73	0.63	0.55	0.48	0.41	0.35	0.29
20			1.87	1.43	1.18	1.01	0.87	0.76	0.66	0.58	0.51	0.44	0.38
21				1.89	1.45	1.21	1.04	0.90	0.79	0.69	0.61	0.54	0.47
22					1.91	1.48	1.24	1.06	0.93	0.82	0.72	0.64	0.57
23						1.93	1.50	1.26	1.09	0.95	0.84	0.75	0.67
24							1.95	1.52	1.28	1.11	0.98	0.87	0.77
25								1.97	1.54	1.30	1.13	1.00	0.89
26									1.98	1.56	1.32	1.15	1.02
27										2.00	1.58	1.34	1.17
28											2.01	1.59	1.36
29												2.03	1.61
30													2.04