3.3: z-score / z-transformation

The z-score is the result of transformation of data that converts a dataset of x values, $\{x_i\}$, that has a mean of $\bar{x}$ and standard deviation s to a set of z values $\{z_i\}$ that has a mean of $\bar{z}=0$ and a standard deviation of $s_z=1$. It will be very useful when we need to compute probabilities associated with normal distributions. The z-transformation is defined by

$$z=\frac{x-\bar{x}}{s}\backslash\backslash\backslash\backslash\backslash(\text{sample}) \tag{3.3.1}$$

$$z=\frac{x-\mu}{\sigma}\backslash\backslash\backslash\backslash\backslash(\text{population}) \tag{3.3.2}$$

Example 3.12 : Find the z-scores of the data given in the left column of the table below.

Data x_i	x_i^2	z-score, z_i
18	324	(18-9.9)/6.2 = 1.3
15	225	(15-9.9)/6.2 = 0.8
12	144	(12-9.9)/6.2 = 0.3
6	36	(6-9.9)/6.2 = -0.6
8	64	(8-9.9)/6.2 = -0.3
2	4	(2-9.9)/6.2 = -1.3
3	9	(3-9.9)/6.2 = -1.1
5	25	(5-9.5)/6.2 = -0.8
20	400	(20-9.5)/6.2 = -1.7
10	100	(10-9.5)/6.2 = 0.1
$\sum x_i = 99$	$\sum x_i^2 = 1331$	

The dataset size is $n=10$. You need to compute the z-score for each data value separately. To do the calculation, both $\bar{x}$ and s are needed. So in addition to the sum of the data, $\sum x$, we also need the sum of the x^2 values. The work of getting those sums is shown in the table above. With the x and x^2 sums we get

$$\bar{x}=\frac{\sum x_i}{n}=\frac{99}{10}=9.9 \tag{3.3.3}$$

and

`s^{2} &=& \frac{\sum x_{i}^{2} - [\frac{(\sum x_i)^2}{n}]}{n-1} =\frac{1331 - [\frac{99^2}{10}]}{9}=\frac{1331-980.1}{9}=39.0`

and $s=\sqrt{39}=6.2$.

Using these values for $\bar{x}$ and s in the third column of the table above, compute the z-scores as shown. If we had computed the z-scores more accurately, they would add up to zero, $\sum z_i = 0$ (the mean of the z-scores is zero.)

□