homework_problem_set_2

JP

9/1/2021

$$s_x^2 = \frac{\Sigma(X - \overline{X})^2}{N - 1}$$

$$s_x = \sqrt{\frac{\Sigma(X - \overline{X})^2}{N - 1}}$$

```
# get the sum
4.377026 + 8.628671 + 6.403844 + 10.191558 + 4.452261 + 7.848496 + 12.25767 + 7.950580 + 9.182271 + 0.7
```

```
## [1] 71.99798
```

```
# sum is 71.99798

# get the mean
71.99798/10
```

```
## [1] 7.199798
```

```
# mean is 7.199798

# get the deviates
4.377026 - 7.199798 # -2.822772
```

```
## [1] -2.822772
```

```
8.628671 - 7.199798 # 1.428873
```

```
## [1] 1.428873
```

```
6.403844 - 7.199798 # -0.795954
```

```
## [1] -0.795954
```

```
10.191558 - 7.199798 # 2.99176
```

```
## [1] 2.99176
```

```r
4.452261 - 7.199798 # -2.747537
```

```
## [1] -2.747537
```

```r
7.848496 - 7.199798 # 0.648698
```

```
## [1] 0.648698
```

```r
12.25767 - 7.199798 # 5.057872
```

```
## [1] 5.057872
```

```r
7.950580 - 7.199798 # 0.750782
```

```
## [1] 0.750782
```

```r
9.182271 - 7.199798 # 1.982473
```

```
## [1] 1.982473
```

```r
0.705607 - 7.199798 # -6.494191
```

```
## [1] -6.494191
```

```r
# square the deviates
(-2.822772)^2
```

```
## [1] 7.968042
```

```r
(1.428873)^2
```

```
## [1] 2.041678
```

```r
(-0.795954)^2
```

```
## [1] 0.6335428
```

```r
(2.99176)^2
```

```
## [1] 8.950628
```

```r
(-2.747537)^2
```

```
## [1] 7.54896
```

```r
(0.648698)^2
```

```
## [1] 0.4208091
```

```r
(5.057872)^2
```

```
## [1] 25.58207
```

```r
(0.750782)^2
```

```
## [1] 0.5636736
```

```r
(1.982473)^2
```

```
## [1] 3.930199
```

```r
(-6.494191)^2
```

```
## [1] 42.17452
```

```r
# get the sum of the squared deviates
7.968042 + 2.041678 + 0.6335428 + 8.950628 + 7.54896 + 0.4208091 + 25.58207 + 0.5636736 + 3.930199 + 42
```

```
## [1] 99.81412
```

```r
# numerator is 99.81412

# denominator is 10 - 1
10-1
```

```
## [1] 9
```

```r
# denominator is 9

# get the variance
99.81412/9
```

```
## [1] 11.09046
```

```r
# variance is 11.09046

# standard deviation is the variance squared
sqrt(11.09046)
```

```
## [1] 3.330234
```

```r
# sd is 3.330234

# double check our work

sd(numbers)
```

```
## [1] 3.330235
```

```r
# pretty close
```

$$s_x^2 = \frac{\Sigma X^2 - \frac{(\Sigma X)^2}{N}}{N-1}$$

$$s_x = \sqrt{\frac{\Sigma X^2 - \frac{(\Sigma X)^2}{N}}{N-1}}$$

```r
numbers
```

```
##  [1]  4.377026  8.628671  6.403844 10.191558  4.452261  7.848496 12.257677
##  [8]  7.950580  9.182271  0.705607
```

```r
4.377026 + 8.628671 + 6.403844 + 10.191558 + 4.452261 + 7.848496 + 12.25767 + 7.950580 + 9.182271 + 0.7
```

```
## [1] 71.99798
```

```r
# sum is 71.99798

71.99798^2
```

```
## [1] 5183.709
```

```r
# squared sum of x is 5183.709

5183.709/10
```

```
## [1] 518.3709
```

```r
# 518.3709 is the value after dividing by N

4.377026^2 + 8.628671^2 + 6.403844^2 + 10.191558^2 + 4.452261^2 + 7.848496^2 + 12.25767^2 + 7.950580^2 +
```

```
## [1] 618.1851
```

```r
# sum of squared Xs is 618.1851

618.1851 - 518.3709
```

```
## [1] 99.8142
```

```
# numerator is 99.8142

# denominator is N - 1
10-1
```

```
## [1] 9
```

```
99.8142/9
```

```
## [1] 11.09047
```

```
# variance is 11.09047

sqrt(11.09047)
```

```
## [1] 3.330236
```

```
# sd is 3.330236

sd(numbers)
```

```
## [1] 3.330235
```

Range

```
data_example <- tibble(Maximum = c(81, 41, 80, 71, 100, 4, 39, 71, 60, 38),
                       Minimum = c(50, 2, 34, 10, 9, 0, 12, 5, 50, 3))

data_example
```

```
## # A tibble: 10 x 2
##    Maximum Minimum
##      <dbl>   <dbl>
##  1      81      50
##  2      41       2
##  3      80      34
##  4      71      10
##  5     100       9
##  6       4       0
##  7      39      12
##  8      71       5
##  9      60      50
## 10      38       3
```

z-Score

$$z = \frac{X - \overline{X}}{S_X}$$

```
get_z <- tibble(single_obs = c(8, 10, 23, 23, 30, 53, 40, 57, 27, 8),
                mean = c(10, 15, 17, 24, 34, 45, 36, 50, 41, 4),
                sd = c(2.1, 3.1, 1.4, 3, 1.2, .2, .9, .47, .16, .67))

get_z
```

```
## # A tibble: 10 x 3
##    single_obs  mean    sd
##         <dbl> <dbl> <dbl>
##  1          8    10  2.1
##  2         10    15  3.1
##  3         23    17  1.4
##  4         23    24  3
##  5         30    34  1.2
##  6         53    45  0.2
##  7         40    36  0.9
##  8         57    50  0.47
##  9         27    41  0.16
## 10          8     4  0.67
```

Raw Score From z-Score

$$X = (z)(S_X) + \overline{X}$$

```
get_raw <- tibble(z = c(1, 2, 3, -3, -2, -1, 1.4, 1.6, -2.4, -3),
                mean = c(10, 15, 17, 24, 34, 45, 36, 50, 41, 4),
                sd = c(2.1, 3.1, 1.4, 3, 1.2, .2, .9, .47, .16, .67))

get_raw
```

```
## # A tibble: 10 x 3
##        z  mean    sd
##    <dbl> <dbl> <dbl>
##  1   1      10  2.1
##  2   2      15  3.1
##  3   3      17  1.4
##  4  -3      24  3
##  5  -2      34  1.2
##  6  -1      45  0.2
##  7   1.4    36  0.9
##  8   1.6    50  0.47
##  9  -2.4    41  0.16
## 10  -3       4  0.67
```

Standard Error of the Mean

$$\sigma_{\overline{X}} = \frac{\sigma_X}{\sqrt{N}}$$

```
get_sample_distribution <- tibble(ind_mean = c(10, 15, 17, 24, 34, 45, 36, 50, 41, 4),
                                  mu = c(12, 16, 14, 22, 30, 40, 32, 54, 43, 8),
                                  sigma_x = c(3, 2, 4, 1.4, .2, .6, .7, .1, 2.4, 2.9), #this is not sig
                                  N = c(10, 12, 14, 16, 20, 24, 11, 8, 15, 10))

get_sample_distribution
```

```
## # A tibble: 10 x 4
##    ind_mean    mu sigma_x     N
##       <dbl> <dbl>   <dbl> <dbl>
##  1       10    12     3      10
##  2       15    16     2      12
##  3       17    14     4      14
##  4       24    22     1.4    16
##  5       34    30     0.2    20
##  6       45    40     0.6    24
##  7       36    32     0.7    11
##  8       50    54     0.1     8
##  9       41    43     2.4    15
## 10        4     8     2.9    10
```

$$z = \frac{\overline{X} - \mu}{\sigma_{\overline{X}}}$$

```
get_sample_distribution
```

```
## # A tibble: 10 x 4
##    ind_mean    mu sigma_x     N
##       <dbl> <dbl>   <dbl> <dbl>
##  1       10    12     3      10
##  2       15    16     2      12
##  3       17    14     4      14
##  4       24    22     1.4    16
##  5       34    30     0.2    20
##  6       45    40     0.6    24
##  7       36    32     0.7    11
##  8       50    54     0.1     8
##  9       41    43     2.4    15
## 10        4     8     2.9    10
```

Let's Try Variance & SD Together Now

```r
set.seed(090221)

numbers = rnorm(n = 10, mean = 4.14, sd = 1.11)

N = 10

numbers
```

```
##  [1] 5.060850 3.865486 3.521444 3.571518 4.629313 4.115526 4.492806 4.364263
##  [9] 5.067677 2.977059
```