**Answer on Question #82458 – Math – Statistics and Probability**

**Question**

Sample of 40 electric batteries gives a mean life span of 600 hours with a standard deviation of 20 hours. Another sample of 50 electric batteries gives a mean lifespan of 520 hours with a standard deviation of 30 hours. If these two samples were combined and used in a given project simultaneously, determine the combined new mean for the larger sample and hence determine the combined or pulled standard deviation.

**Solution**

Let $m$ be the combined mean.
Let $x_1$ be the mean of first sample.
Let $x_2$ be the mean of the second sample.
Let $n_1$ be the size of the first sample.
Let $n_2$ be the size of the second sample.
Let $s_1$ be the standard deviation of the first sample.
Let $s_2$ be the standard deviation of the second sample.

$$\text{combined mean} = m = \frac{n_1 x_1 + n_2 x_2}{n_1 + n_2}$$

$$\text{combined standard deviation} = \sqrt{\frac{n_1 s_1^2 + n_1(m - x_1)^2 + n_2 s_2^2 + n_2(m - x_2)^2}{n_1 + n_2}}$$

$$s^2 = \frac{n_1 s_1^2 + n_1(m - x_1)^2 + n_2 s_2^2 + n_2(m - x_2)^2}{n_1 + n_2} =$$

$$= \frac{\sum_{i=1}^{n_1}(x_i - x_1)^2 + n_1(m - x_1)^2 + \sum_{i=1}^{n_2}(y_i - x_2)^2 + n_2(m - x_2)^2}{n_1 + n_2} =$$

$$= \frac{\sum_{i=1}^{n_1}((x_i - x_1)^2 + (x_1 - m)^2) + \sum_{i=1}^{n_2}((y_i - x_2)^2 + (x_2 - m)^2)}{n_1 + n_2} =$$

$$= \frac{\sum_{i=1}^{n_1}({x_i}^2 - 2x_i x_1 + 2{x_1}^2 - 2m x_1 + m^2)}{n_1 + n_2} +$$

$$+ \frac{\sum_{i=1}^{n_2}({y_i}^2 - 2y_i x_2 + 2{x_2}^2 - 2m x_2 + m^2)}{n_1 + n_2} =$$

$$= \frac{\sum_{i=1}^{n_1}\left({x_i}^2 + m^2 - 2m\sum_{j=1}^{n_1}\frac{x_j}{n_1}\right) + 2n_1{x_1}^2 - 2x_1\sum_{i=1}^{n_1} x_i}{n_1 + n_2} +$$

$$+ \frac{\sum_{i=1}^{n_2}\left({y_i}^2 + m^2 - 2m\sum_{j=1}^{n_2}\frac{y_j}{n_2}\right) + 2n_2{x_2}^2 - 2x_2\sum_{i=1}^{n_2} y_i}{n_1 + n_2} =$$

$$=\frac{\sum_{i=1}^{n_1}\left(x_i{}^2+m^2-2m\sum_{j=1}^{n_1}\frac{x_j}{n_1}\right)+2n_1x_1{}^2-2x_1n_1x_1}{n_1+n_2}+$$

$$+\frac{\sum_{i=1}^{n_2}\left(y_i{}^2+m^2-2m\sum_{j=1}^{n_2}\frac{y_j}{n_2}\right)+2n_2x_2{}^2-2x_2n_2x_2}{n_1+n_2}=$$

$$=\frac{\sum_{i=1}^{n_1}\left(x_i{}^2+m^2-2m\sum_{j=1}^{n_1}\frac{x_j}{n_1}\right)}{n_1+n_2}+\frac{\sum_{i=1}^{n_2}\left(y_i{}^2+m^2-2m\sum_{j=1}^{n_2}\frac{y_j}{n_2}\right)}{n_1+n_2}$$

Since each $\left(-2m\sum_{j=1}^{n_1}\frac{x_j}{n_1}\right)$ term appears $n_1$ times , then we can reorder the sums

$$\sum_{i=1}^{n_1}\left(x_i{}^2+m^2-2m\sum_{j=1}^{n_1}\frac{x_j}{n_1}\right)=\sum_{i=1}^{n_1}(x_i{}^2+m^2-2mx_i)$$

Hence,

$$s^2=\frac{n_1s_1^2+n_1(m-x_1)^2+n_2s_2^2+n_2(m-x_2)^2}{n_1+n_2}=$$

$$=\frac{\sum_{i=1}^{n_1}(x_i{}^2+m^2-2mx_i)}{n_1+n_2}+\frac{\sum_{i=1}^{n_2}(y_i{}^2+m^2-2my_i)}{n_1+n_2}=$$

$$=\frac{\sum_{i=1}^{n_1+n_2}(z_i{}^2+m^2-2mz_i)}{n_1+n_2}=$$

$$=\frac{\sum_{i=1}^{n_1+n_2}(z_i-m)^2}{n_1+n_2}\stackrel{\text{def}}{=}s^2$$

$n_1=40, x_1=600\ h, s_1=20\ h$
$n_2=50, x_2=520\ h, s_2=30\ h$

$$\text{combined mean}=m=\frac{40(600)+50(520)}{40+50}=\frac{5000}{9}\approx 555.56\ (h)$$

combined standard deviation =

$$=\sqrt{\frac{40(20)^2+40\left(\frac{5000}{9}-600\right)^2+50(30)^2+50\left(\frac{5000}{9}-520\right)^2}{40+50}}\approx 47.52\ (h)$$

combined mean = 555.56 hours

combined standard deviation = 47.52 hours