Your Mean May Not Mean What You Mean It to Mean.

Kenneth L. (Ken) Johnson
NASA Engineering and Safety Center (NESC)
April 10, 2019/ DATAWorks 2019
POP QUIZ.

• What is the average of apple and orange?
Your presenter, Ken Johnson.

- Statistician, NASA Engineering and Safety Center (NESC) [https://www.nasa.gov/nesc](https://www.nasa.gov/nesc).
- Lead, NASA Statistical Engineering Team.
- Industrial quality, management, tech service, sales and other pursuits.
  - Aluminum/steel coil coating.
  - Polymer coatings.
  - Rigid urethane foam.
- MS Operations Research/Applied Statistics and Quality, University of Alabama in Huntsville.
- (Nearly an) MBA, DePaul University, Chicago.
- BA Chemistry, Grinnell College.
Objective of this presentation.

- **Success:** you leave this room convinced that calculating a mean is different from calculating a mean that means something.
What you’ll hear here.

• How to calculate a mean.
• How to calculate a mean badly.
• Why it’s bad to do that.
• How to do it better.
• How to do it even better than that.
• When to call in the statistician.
• The wrap.
How to Calculate a Mean.
How to calculate a mean.

Geeks.

- Given a list of values with \( n \) elements, the mean is defined to be
  \[
  \mu(x) = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}.
  \]

Actual People.

- Add up your list of numbers.
- Divide that by the number of numbers in your list.
- Done.
How to calculate a mean, according to statisticians.

• It depends on your question and the data.
How to Calculate a Mean Badly.
How to calculate a mean badly.

• Add up your list of numbers.
• Divide that by the number of numbers in your list.
• Done.

• But what if your list contains non-random structure?
Your experiment.

- Problem statement:
  - Whose battery stays cooler when it’s charged over the range of use temperatures: Brand X or Y Corporation?

- Response: Temperature after charging, $F$

- Fixed (nonrandom, structural) factors:
  - Battery Brand X 2 [Brand X, Y Corp]
  - Environmental Temperature [RT, 140] $F$

<table>
<thead>
<tr>
<th>Test No</th>
<th>Batt Brand</th>
<th>Env Temp F</th>
<th>End Temp F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Brand X</td>
<td>-40</td>
<td>-6</td>
</tr>
<tr>
<td>2</td>
<td>Brand X</td>
<td>Room</td>
<td>60</td>
</tr>
<tr>
<td>3</td>
<td>Brand X</td>
<td>Room</td>
<td>56</td>
</tr>
<tr>
<td>4</td>
<td>Brand X</td>
<td>Room</td>
<td>62</td>
</tr>
<tr>
<td>5</td>
<td>Brand X</td>
<td>140</td>
<td>172</td>
</tr>
<tr>
<td>6</td>
<td>Brand X</td>
<td>140</td>
<td>186</td>
</tr>
<tr>
<td>7</td>
<td>Y Corp</td>
<td>Room</td>
<td>50</td>
</tr>
<tr>
<td>8</td>
<td>Y Corp</td>
<td>Room</td>
<td>55</td>
</tr>
<tr>
<td>9</td>
<td>Y Corp</td>
<td>Room</td>
<td>57</td>
</tr>
<tr>
<td>AVG</td>
<td></td>
<td></td>
<td>75.9</td>
</tr>
</tbody>
</table>

**POP QUIZ**
What does this average mean?
What does this average mean?

- Nothing.
- No, really. Nothing.
  - Nada.
  - Zilch.
  - Zip.
  - Zero.
  - **Nothing.**

- Just because you can calculate it doesn’t mean it answers your question ...
- … **Or even makes sense.**

### Data Table

<table>
<thead>
<tr>
<th>Test No</th>
<th>Batt Brand</th>
<th>Env Temp F</th>
<th>End Temp F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Brand X</td>
<td>-40</td>
<td>-6</td>
</tr>
<tr>
<td>2</td>
<td>Brand X</td>
<td>Room</td>
<td>60</td>
</tr>
<tr>
<td>3</td>
<td>Brand X</td>
<td>Room</td>
<td>56</td>
</tr>
<tr>
<td>4</td>
<td>Brand X</td>
<td>Room</td>
<td>62</td>
</tr>
<tr>
<td>5</td>
<td>Brand X</td>
<td>140</td>
<td>172</td>
</tr>
<tr>
<td>6</td>
<td>Brand X</td>
<td>140</td>
<td>186</td>
</tr>
<tr>
<td>7</td>
<td>Y Corp</td>
<td>Room</td>
<td>50</td>
</tr>
<tr>
<td>8</td>
<td>Y Corp</td>
<td>Room</td>
<td>55</td>
</tr>
<tr>
<td>9</td>
<td>Y Corp</td>
<td>Room</td>
<td>57</td>
</tr>
<tr>
<td><strong>AVG</strong></td>
<td></td>
<td><strong>75.9</strong></td>
<td></td>
</tr>
</tbody>
</table>

**Problem statement:**
What’s the average of Brand X and Y Corp?
Structure isn’t random.

- The day you fly you will have chosen either a Brand X or a Y Corporation battery.
  - Factor Batt Brand is driven by choice, not randomness.
  - Makes no sense to talk about an average of Brand X and Y Corporation.

- You won’t know which individual battery you will launch: random choice.
  - The average of the sample of Brand X batteries looks through the fog of uncertainty to get the best estimate of the Brand X battery you will fly.
How to Calculate a Mean Better.
Problem statement:
- Whose battery stays cooler when it’s charged: Brand X or Y Corporation?

Problem statement, stated so the problem can be solved using the data (statistics):
- Compare averages: Brand X vs Y Corporation.

Details.
- Make sure they’re at the same temperature!
  - Temperature is also a structural variable.
- Make sure the difference in averages isn’t just by chance.
  - Significance

<table>
<thead>
<tr>
<th>Test No</th>
<th>Batt Brand</th>
<th>Env Temp F</th>
<th>End Temp F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Brand X</td>
<td>-40</td>
<td>-6</td>
</tr>
<tr>
<td>2</td>
<td>Brand X</td>
<td>Room</td>
<td>60</td>
</tr>
<tr>
<td>3</td>
<td>Brand X</td>
<td>Room</td>
<td>56</td>
</tr>
<tr>
<td>4</td>
<td>Brand X</td>
<td>Room</td>
<td>62</td>
</tr>
<tr>
<td>5</td>
<td>Brand X</td>
<td>140</td>
<td>172</td>
</tr>
<tr>
<td>6</td>
<td>Brand X</td>
<td>140</td>
<td>186</td>
</tr>
<tr>
<td>7</td>
<td>Y Corp</td>
<td>Room</td>
<td>50</td>
</tr>
<tr>
<td>8</td>
<td>Y Corp</td>
<td>Room</td>
<td>55</td>
</tr>
<tr>
<td>9</td>
<td>Y Corp</td>
<td>Room</td>
<td>57</td>
</tr>
</tbody>
</table>

Analysis: compare means of 2 groups of sample size = 3.
Comparison of means for END TEMP.

• Mean of Brand X: 59.3 +/- 7.59
  • 95.0% confidence interval [51.74, 66.93])

• Mean of Y Corp: 54.0 +/- 8.96
  • 95.0% confidence interval [45.04, 62.96]

• Difference between the means = 5.3 +/- 7.58
  • 95.0% confidence interval assuming equal variances:
    [-2.24, 12.91]

  **English:** NO DIFFERENCE WAS DETECTED
  • The difference might be zero …
  • … the Y Corp battery might even be hotter!

• t test to compare means
  • Null hypothesis: mean1 = mean2
  • Alt. hypothesis: mean1 NE mean2
  • assuming equal variances: t = 1.95471  **P-value = 0.13**

  • Do not reject the null hypothesis for α = 0.05.

*This is your TRUTH.*
Is 59.3 The Mean for ALL Brand X batteries?

NO, but it’s your best estimate.

• Mean of Brand X: 59.3 ± 7.59

The next experiment will give a different mean – though it should be CLOSE.

It isn’t The Real True Mean unless you’ve measured everything.

Most of the time, this interval should capture the true unknown mean.
Is 88.3 the best estimate mean for Brand X batteries?

If nonrandom factors are present ...

Mean of Brand X: 88.3 +/- ????

... you might as well be averaging apples and oranges!

The uncertainty has to include the probability of being at either RT vs -40 F vs 140 F.
Even Better: Plan.
Plan your experiment to be efficient.

- Problem statement:
  - Which battery stays cooler when it’s charged over the range of use temperatures: Brand X or Y Corporation?

- Problem statement, stated so the problem can be solved using the data (statistics):
  - Evaluate the significance of the difference in averages: Brand X vs Y Corporation.

- Test matrix details.
  - Balanced and orthogonal.
  - Randomized.
  - Quantified factors.
  - Planned using DOE tools to help right-size the experiment.

<table>
<thead>
<tr>
<th>Test No</th>
<th>Batt Brand</th>
<th>Env Temp F</th>
<th>End Temp F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Y Corp</td>
<td>70</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>Y Corp</td>
<td>140</td>
<td>162</td>
</tr>
<tr>
<td>3</td>
<td>Brand X</td>
<td>74</td>
<td>62</td>
</tr>
<tr>
<td>4</td>
<td>Brand X</td>
<td>149</td>
<td>186</td>
</tr>
<tr>
<td>5</td>
<td>Brand X</td>
<td>70</td>
<td>60</td>
</tr>
<tr>
<td>6</td>
<td>Y Corp</td>
<td>79</td>
<td>57</td>
</tr>
<tr>
<td>7</td>
<td>Brand X</td>
<td>136</td>
<td>172</td>
</tr>
<tr>
<td>8</td>
<td>Y Corp</td>
<td>140</td>
<td>168</td>
</tr>
</tbody>
</table>

Finding the difference in means of 2 groups of sample size = 4 using linear regression to control for Environmental Temperature.
Comparison of means for **END TEMP**.

**END TEMP** = -65.1859 + 5.375*Brand + 1.67656*EnvTemp

*where*

**Brand** = 1 if Batt Brand = Brand X,
-1 if Batt Brand = Y Corp, 0 otherwise

<table>
<thead>
<tr>
<th>Batt Brand</th>
<th>Count</th>
<th>LS Mean</th>
<th>LS Sigma</th>
<th>Homogeneous Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y Corp</td>
<td>4</td>
<td>109.25</td>
<td>2.18</td>
<td>X</td>
</tr>
<tr>
<td>Brand X</td>
<td>4</td>
<td>120.00</td>
<td>2.18</td>
<td>X</td>
</tr>
</tbody>
</table>

**Method**: 95.0 percent LSD

- **Difference between the means** = 10.75 F +/- 7.93

*Statistically significant difference at p < 0.05.*

**English**: The difference is **STATISTICALLY SIGNIFICANT**.
- The difference is greater than zero ...  
- ... the Y Corp battery clearly charges cooler!

---

3/8/2019 Mean^4 for DATAWorks 2019
When to call a statistician.

- If any part of this puzzles you.
  - Even if it makes sense – an experimental design SME may help your team get maximum value from test and the data.

- First – **before you plan** the experiment.
Bottom lines.

• Means mean something. Calculate and interpret them with care.
  • Don’t just do math.
• Start with the problem statement.
• Design your experiment to address your problem statement efficiently.
• Call an SME when you need an SME.
  • That SME may be a statistician.

Thank you!
Questions?