Using Hands-on Methods with Computer Simulations to Teach Sampling Distributions and Inference

Amy G. Froelich
William M. Duckworth
W. Robert Stephenson

Department of Statistics
Iowa State University

Abstract

In the past, we have used hands-on activities coupled with Java scripts on the web to teach students about sampling distributions and inference. However, students had difficulties connecting the hands-on activities with the java scripts. The hands-on activities are now duplicated on the computer using JMP scripts.

Goal

Using the hands-on activities and the JMP scripts, students will discover concepts related to sampling distributions and inference such as

- What is a sampling distribution?
- What does C% confident really mean?
- What is Type I error?
- What is Power?

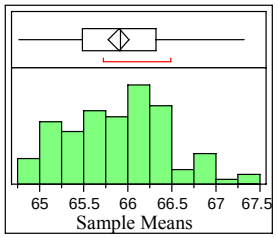
What is a sampling distribution?

Hands-on Activity

- Population of 250 values.
- Each student samples from population by hand using random number table.
  - n = 1
  - n = 5
  - n = 10
- Calculate sample mean.
- Combine sample means from class into histogram.

Computer Simulation

- Population
- JMP Script samples from population 100 times with n = 25.
- Records 100 sample mean values.
- Histogram of results.



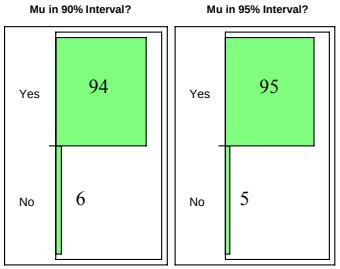
What does C% confidence really mean?

Hands-on Method

- Population of 250 values.
- Each student samples from population by hand using random number table with n = 10
- Calculate sample mean and confidence interval.
- Is μ in your confidence interval?
- How many people have μ in their confidence intervals?

Computer Simulation

- Population
- JMP Script samples from population 100 times with n = 25.
- Records 100 sample mean values and 100 confidence intervals.
- How many 90% and 95% confidence intervals contain μ.



What is Type I Error?

Hands-on Method

- Population of 250 values.
- Each student samples from population by hand using random number table with n = 10
- Calculate sample mean and z test statistic using true value of μ.
- Calculate p-value.
- Will you reject H_o at α = 0.1, 0.05, or 0.01?
- How many people will reject H_o?

Computer Simulation

- Population
- JMP Script samples from population 100 times with n = 25.
- Records 100 sample mean values, 100 z test statistics, and 100 p-values.
- How many times will you reject H_o?

Percent of Rejections of H_o

$\alpha = 0.1$	11%
$\alpha = 0.05$	5%
$\alpha = 0.01$	1%

What is power?

Problem: Two varieties of corn are planted in 36 plots in the same field. Do Type A and Type B produce different mean corn yields?

Convenience Assignment

A 130	A 149	A 139	B 155	B 137	B 145
A 149	A 133	A 152	B 131	B 147	B 136
A 141	A 156	A 137	B 146	B 132	B 148
A 150	A 142	A 155	B 136	B 152	B 133
A 139	A 155	A 139	B 147	B 137	B 153
A 155	A 138	A 150	B 137	B 145	B 136

Conclusion: The mean yield of Type A is slightly higher than Type B but the difference is not statistically significant.

Systematic Assignment

A 130	B 137	A 139	B 155	A 149	B 145
B 137	A 133	B 140	A 143	B 147	A 148
A 141	B 144	A 137	B 146	A 144	B 148
B 138	A 142	B 143	A 148	B 152	A 145
A 139	B 143	A 139	B 147	A 149	B 153
B 143	A 138	B 138	A 149	B 145	A 148

Conclusion: The mean yield of Type B is now slightly higher than Type A, but the difference is not statistically significant.

Why do two assignments give different results? What kind of assignment should be used?

Hands-on Activity

Students make a random assignment of the two varieties to the 36 plots using a 6-sided die. Using the TRUTH, students assign values to each of the 36 randomly assigned plots and conduct a hypothesis test to determine if Type A has a better mean corn yield than Type B.

THE TRUTH

A=130 B=118	A=149 B=137	A=139 B=127	A=167 B=155	A=149 B=137	A=157 B=145
A=149 B=137	A=133 B=121	A=152 B=140	A=143 B=131	A=159 B=147	A=148 B=136
A=141 B=129	A=156 B=144	A=137 B=125	A=158 B=146	A=144 B=132	A=160 B=148
A=150 B=138	A=142 B=130	A=155 B=143	A=148 B=136	A=164 B=152	A=145 B=133
A=139 B=127	A=155 B=143	A=139 B=127	A=159 B=147	A=149 B=137	A=165 B=153
A=155 B=143	A=138 B=126	A=150 B=138	A=149 B=137	A=157 B=145	A=148 B=136

A Random Assignment

B 118	A 149	A 139	A 167	A 149	A 157
B 137	B 121	B 140	A 143	B 147	B 136
B 129	B 144	A 137	B 146	B 132	A 160
B 138	A 142	A 155	B 136	A 164	A 145
A 139	A 155	A 139	A 159	B 137	A 165
B 143	B 126	B 138	B 137	B 145	A 148

Under the random assignment, there is an overwhelming statistically significant difference in mean corn yields between Type A and B, with Type A having higher mean yields.

Computer Simulation

- JMP script makes 100 random assignments of varieties A and B to the 36 plots.
- Complete test of
  - Ho: $\mu_A - \mu_B = 0$
  - Ha: $\mu_A - \mu_B \neq 0$
- Null hypothesis is incorrect ($\mu_A - \mu_B = 12$)
- How many times will you reject H_o?

Further Simulation

- Change true value of $\mu_A - \mu_B$
- Null hypothesis is still incorrect.
- How many times will you reject H_o?
- How does this relate to α?

Percent of Rejections of H_o

	$\mu_A - \mu_B = 3$	$\mu_A - \mu_B = 6$	$\mu_A - \mu_B = 12$
$\alpha = 0.05$	11%	46%	96%
$\alpha = 0.01$	1%	17%	88%