Getting Started

Andrew Marshall

5/30/2018

# Instructions

This document runs a simple analysis of the Table 1 from (Wansink and Payne 2009). Edit the header information to show your name and the date you complete the assignment.

Modify this document to analyze either Calories per Serving or Servings per Recipe. Document any changes you make in the literate portion of the file. Comment on your choice of measure to analyze.

Change the name of this file to match your user name on D2L, keeping the ‘Rmd’ extension, and include week number in the title (for example, Peter.Claussen.1.Rmd). Upload this file to D2L. Typeset this file to Word or PDF and upload the result to D2L as well.

## Data

Mean and (SD) for selected recipes from “Joy of Cooking”

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Measure | 1936 | 1946 | 1951 | 1963 | 1975 | 1997 | 2006 |
| calories per recipe (SD) | 2123.8 (1050.0) | 2122.3 (1002.3) | 2089.9 (1009.6) | 2250.0 (1078.6) | 2234.2 (1089.2) | 2249.6 (1094.8) | 3051.9 (1496.2) |
| calories per serving (SD) | 268.1 (124.8) | 271.1 (124.2) | 280.9 (116.2) | 294.7 (117.7) | 285.6 (118.3) | 288.6 (122.0) | 384.4 (168.3) |
| servings per recipe (SD) | 12.9 (13.3) | 12.9 (13.3) | 13.0 (14.5) | 12.7 (14.6) | 12.4 (14.3) | 12.4 (14.3) | 12.7 (13.0) |

# Analysis

### Enter data

CookingTooMuch.dat <- data.frame(  
 Year=c(1936, 1946, 1951, 1963, 1975, 1997, 2006),  
 CaloriesPerRecipeMean = c(2123.8, 2122.3, 2089.9, 2250.0, 2234.2, 2249.6, 3051.9),  
 CaloriesPerRecipeSD = c(1050.0, 1002.3, 1009.6, 1078.6, 1089.2, 1094.8, 1496.2),  
 CaloriesPerServingMean = c(268.1, 271.1, 280.9, 294.7, 285.6, 288.6, 384.4),  
 CaloriesPerServingSD = c(124.8, 124.2, 116.2, 117.7, 118.3, 122.0, 168.3),  
 ServingsPerRecipeMean = c(12.9, 12.9, 13.0, 12.7, 12.4, 12.4, 12.7),  
 ServingsPerRecipeSD = c(13.3, 13.3, 14.5, 14.6, 14.3, 14.3, 13.0)  
)

### Create values for confidence interval plot

Wansink reports that 18 recipes were analyzed.

n <- 18

Assume a significance level of 5%.

alpha <- 0.05

Use standard formula for standard error and confidence interval .

StandardError <- function(sigma, n) {  
 sigma/sqrt(n)  
}  
ConfidenceInterval <- function(sigma, n) {  
 qt(1-alpha/2, Inf)\*StandardError(sigma,n)  
}

Create a variable for plotting and calculate upper and lower bounds using confidence intervals.

PlotCookingTooMuch.dat <- CookingTooMuch.dat  
PlotCookingTooMuch.dat$CaloriesPerServing <- PlotCookingTooMuch.dat$CaloriesPerServingMean  
PlotCookingTooMuch.dat$Lower <- PlotCookingTooMuch.dat$CaloriesPerServing - ConfidenceInterval(CookingTooMuch.dat$CaloriesPerServingSD,n)  
PlotCookingTooMuch.dat$Upper <- PlotCookingTooMuch.dat$CaloriesPerServing + ConfidenceInterval(CookingTooMuch.dat$CaloriesPerServingSD,n)  
PlotCookingTooMuch.dat <- PlotCookingTooMuch.dat[,c("Year","CaloriesPerServing","Lower","Upper")]

Examine the values to make sure we’ve entered correctly.

print(PlotCookingTooMuch.dat)

## Year CaloriesPerServing Lower Upper  
## 1 1936 268.1 210.4464 325.7536  
## 2 1946 271.1 213.7236 328.4764  
## 3 1951 280.9 227.2193 334.5807  
## 4 1963 294.7 240.3264 349.0736  
## 5 1975 285.6 230.9492 340.2508  
## 6 1997 288.6 232.2399 344.9601  
## 7 2006 384.4 306.6508 462.1492

Check our calculations. Wansink reports 95% confidence intervals for the 1936 and 2006 means as [210.4, 325.8] and [359.1, 514.7], respectively. We should be equal, to within rounding error.

CompValues <- PlotCookingTooMuch.dat[c(1,7),c("Lower","Upper")]  
ReferenceValues <- matrix(c(210.4, 325.8, 359.1, 514.7),nrow=2,byrow=TRUE)  
CompValues

## Lower Upper  
## 1 210.4464 325.7536  
## 7 306.6508 462.1492

ReferenceValues

## [,1] [,2]  
## [1,] 210.4 325.8  
## [2,] 359.1 514.7

any(abs(CompValues-ReferenceValues)>0.1)

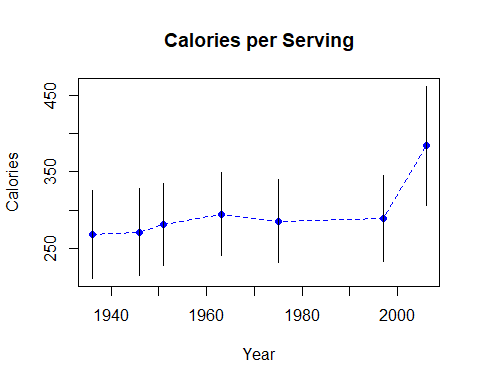
## [1] TRUE

We no longer need the original data.

CookingTooMuch.dat <- NULL

# Plot the table

plot(CaloriesPerServing ~ Year, data=PlotCookingTooMuch.dat,  
 col="blue", pch=19,  
 main="Calories per Serving",   
 ylab="Calories",   
 ylim=c(min(PlotCookingTooMuch.dat$Lower), max(PlotCookingTooMuch.dat$Upper)))  
lines(CaloriesPerServing ~ Year, data=PlotCookingTooMuch.dat,   
 lty="dashed", col="blue", lend=2)  
segments(x0=PlotCookingTooMuch.dat$Year,   
 y0=PlotCookingTooMuch.dat$Lower,   
 x1=PlotCookingTooMuch.dat$Year,   
 y1=PlotCookingTooMuch.dat$Upper)



# Comments

From this plot, it appears that the average calories per serving showed an 35.2% overall increase during the 70 year period. According to Wansink, this increase is mainly due to the use of higher-calorie ingredients and in part to a small increase in serving sizes between 1936 and 1997. However, there is a 33.2% increase shown between 1997 to 2006. Therefore, the recommended comparison should be between the data provided for 1936 and 2006 only.

# References

Wansink, Brian, and Collin Payne. 2009. “The Joy of Cooking Too Much: 70 Years of Calorie Increases in Classic Recipes” 150 (150): 291–92.