

Paired Samples and Blocks (Chapter 21)

Patrick Frenett and Nicholas Horton (nhorton@amherst.edu)

June 16, 2016

Introduction and background

This document is intended to help describe how to undertake analyses introduced as examples in the Fourth Edition of *Intro Stats* (2013) by De Veaux, Velleman, and Bock. More information about the book can be found at http://wps.aw.com/aw_deveaux_stats_series. This file as well as the associated R Markdown reproducible analysis source file used to create it can be found at <http://www.amherst.edu/~nhorton/sdm4>.

This work leverages initiatives undertaken by Project MOSAIC (<http://www.mosaic-web.org>), an NSF-funded effort to improve the teaching of statistics, calculus, science and computing in the undergraduate curriculum. In particular, we utilize the `mosaic` package, which was written to simplify the use of R for introductory statistics courses. A short summary of the R needed to teach introductory statistics can be found in the `mosaic` package vignettes (<http://cran.r-project.org/web/packages/mosaic>).

Note that some of the figures in this document may differ slightly from those in the IS4 book due to small differences in datasets. However in all cases the analysis and techniques in R are accurate.

Chapter 21: Paired samples and Blocks

Section 21.1: Paired data

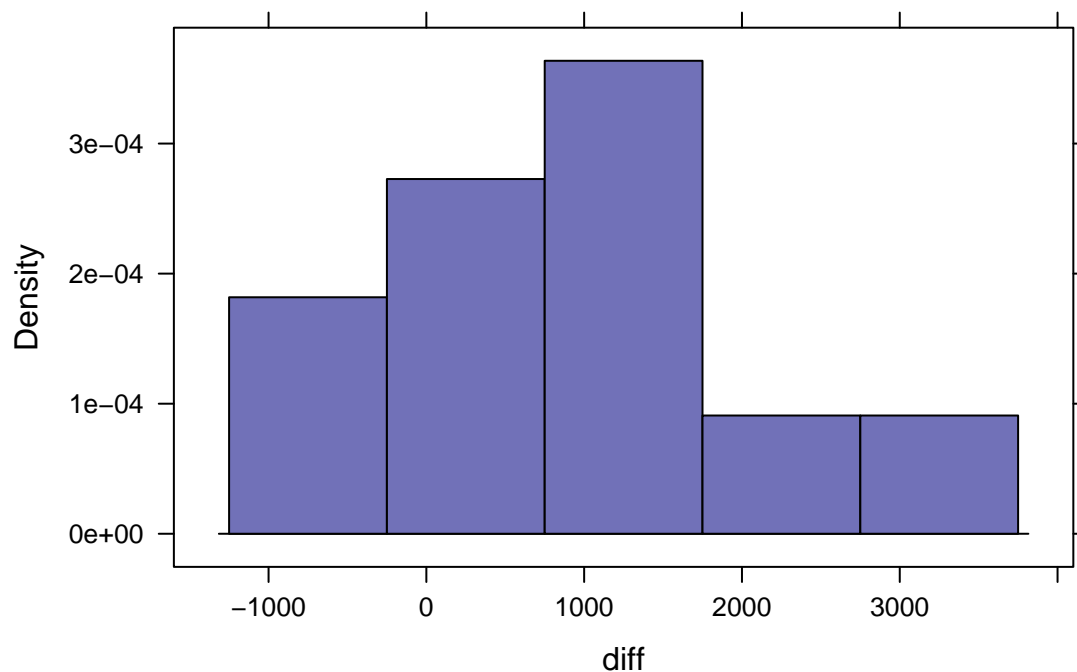
The example on page 586 compares the mileage of 11 field workers using either a 5 day or 4 day schedule.

```
require(mosaic)
fiveday <- c(2798, 7724, 7505, 838, 4592, 8107, 1228, 8718, 1097, 8089, 3807)
fourday <- c(2914, 6112, 6177, 1102, 3281, 4997, 1695, 6606, 1063, 6392, 3362)
ds <- data.frame(fiveday, fourday)
ds <- mutate(ds, diff = fiveday - fourday); ds
```

##	fiveday	fourday	diff
## 1	2798	2914	-116
## 2	7724	6112	1612
## 3	7505	6177	1328
## 4	838	1102	-264
## 5	4592	3281	1311
## 6	8107	4997	3110
## 7	1228	1695	-467
## 8	8718	6606	2112
## 9	1097	1063	34
## 10	8089	6392	1697
## 11	3807	3362	445

Section 21.2: Assumptions and conditions

```
histogram(~ diff, width=1000, center=500/2, data=ds) # page 589
```



```
t.test(~ diff, data=ds)
```

```
##
## One Sample t-test
##
## data: ds$diff
## t = 2.858, df = 10, p-value = 0.01701
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  216.4276 1747.5724
## sample estimates:
## mean of x
##      982
```

Section 21.3: Confidence intervals for matched pairs

The same result is seen as on page 595 for the confidence interval for the population difference in mileage using the (results not shown).

```
t.test(~ diff, data=ds)$conf.int
```