---

title: "Bayesian Single-Case Analysis for AB Designs"

author: "Compiled by Brandon K. Schultz, Ed.D., NCSP"

date: "January 15, 2016"

output: word\_document

---

```{r data\_entry, echo=FALSE}

# Enter your data in the "scdata" line below, between the parentheses, placing a comma between each measurement occasion. The same metric (e.g., 0-100 scale) must be used at each occasion, and the lag between measurement occasions must be consistent. If planned measurement occasions were missed, enter NA in those positions in order to keep the schedule as consistent as possible.

scdata <- c(87.5, NA, 53.4, 72.3, 94.2, 96.6, 57.4, 78.1, 47.2, NA, 82.1, 73.7, 49.3, 79.3, NA, 57.3, 31.7, 50.4, 77.8, 67, 40.5, 1.6, NA, 3.2, 24.1)

n1 = 10 # This is the number of data points that were collected at baseline (A phase)

n2 = 15 # This is the number of data points collected during/after intervention (B phase)

lonum = min(scdata)

hinum = max(scdata)

ypre = scdata[1:n1]

ypost = scdata[n1 + 1:n2]

ssd1 <- data.frame(myy = ypre, myphase = "A")

ssd2 <- data.frame(myy = ypost, myphase = "B")

ssdall = rbind(ssd1, ssd2)

ssdall["xrownum"] <- c(1:nrow(ssdall))

```

Below is an analysis of AB (two-phase) single-case data using a Bayesian approach described by de Vries and Morey (2013). For each analysis, there is a matching graph to illustrate the analyses, as well as notations to help interpret the results.

First, we begin with a simple line graph of the data, with a vertical line separating the baseline and treatment data. Make sure that the data in the graph are accurate before proceding.

#####Figure 1: Overview of Single-Case Data

```{r plot\_data, dpi=300, echo=FALSE, warning=FALSE}

suppressPackageStartupMessages(library(SSDforR))

suppressPackageStartupMessages(require(SSDforR))

library(SSDforR)

require(SSDforR)

suppressPackageStartupMessages(library(ggplot2))

library(ggplot2)

suppressPackageStartupMessages(library(rmarkdown))

library(rmarkdown)

Trendplot <- ggplot(ssdall, aes(x=xrownum, y=myy, colour=myphase), na.rm=TRUE) +

geom\_vline(linetype=2, xintercept=n1+.5)

Trendplot + geom\_line(size=1) + geom\_point() +

theme(legend.title=element\_blank()) +

labs(x = "Time", y = "Data")

```

# I. Descriptive Statistics

One assumption of the Bayesian tests below is normal distributions in the data. If this assumption is significantly violated, the analysis can be adjusted, but the default settings of the Bayesian tests below are likely to lead to misleading results. Below is the results of Shapiro-Wilks test.

```{r, dpi=300, echo=FALSE}

## Perform the test

shapiro.test(ypre); shapiro.test(ypost)

## Plot using a qqplot

#qqnorm(ypre);qqline(ypre, col = 2)

#qqnorm(ypost);qqline(ypost, col = 2)

```

If either test is significant (\*p\* < .05), the data are mostly likely not suited to the tests below.

A boxplot is a good way to visually summarize possible outliers in the two phases, which could also cause problems in extreme instances.

##### Figure 2: Boxplot of AB Phases

```{r, dpi=300, echo=FALSE, warning=FALSE}

attach(ssdall)

descBoxplot <- ggplot(ssdall, aes(myphase, myy))

descBoxplot + geom\_boxplot(na.rm=TRUE) +

theme(legend.title=element\_blank()) +

labs(x = "Phases", y = "Data")

```

The descriptive statistics, including the mean, median, and SD are provided in Box 1 below. The "10% Trim Mean" removes the influence of extreme scores and, as a result, provides the best summary of the two phases if there are concerns with outliers, trends, or autocorrelation. The interquartile range, or "IQR," provides the range from the 25%ile to the 75%ile in each distribution, as displayed in Figure 2 above, but this can be influenced by unstable, trending data.

##### Box 1: Descriptive Statistics

```{r, echo=FALSE, warning=FALSE}

ABdescrip(myy, myphase)

```

# II. Bayesian Single Case Analysis: Change in Levels

Bayesian statistics offer an alternative that does not make the same assumptions as 'classical,' frequentist statistics (e.g., normal distributions). Below is an application of a modified Bayesian t-test to single subject data that was created by De Vries and Morey (2013). In short, the resulting "Bayes Factor" tests the hypothesis that the overall level of behavior changed, coinciding with the change in phases, while actively adjusting for the autocorrelation found in the model. Once autocorrelation is accounted for, level changes from Phases A to B can be accurately estimated. Here, we assume relative stability in each phase, and we are not concerned with trends in the data.

In Figure 3, a graph is provided to represent the tests below, with means in each phase represented by dashed, horizontal lines.

#### Figure 3: Level Change from Phase A to Phase B

```{r, dpi=300, echo=FALSE, warning=FALSE }

Trendplot <- ggplot(ssdall, aes(x=xrownum, y=myy, colour=myphase)) +

geom\_vline(linetype=2, xintercept=n1+.5)

Trendplot + geom\_line(size=1, na.rm=TRUE) + geom\_point(na.rm=TRUE) +

#scale\_x\_continuous(breaks = c(1,3,5,7,9,11,13,15)) +

theme(legend.title=element\_blank()) +

labs(x = "Time", y = "Data") +

geom\_segment(aes(x=1, y=mean(ssd1$myy, na.rm=TRUE), xend=n1, yend=mean(ssd1$myy, na.rm=TRUE)), linetype="longdash", color="tomato") +

geom\_segment(aes(x=n1+1, y=mean(ssd2$myy, na.rm=TRUE), xend=n1+n2, yend=mean(ssd2$myy, na.rm=TRUE)), linetype="longdash", color="turquoise3")

```

```{r, echo=FALSE, results="hide"}

library(BayesSingleSub)

library(coda)

output.JZS.AR = ttest.Gibbs.AR(ypre, ypost, iterations = 20000, return.chains = TRUE, progress=FALSE)

chains.JZS.AR = output.JZS.AR$chains

```

We begin by estimating autocorrelation. The histogram in Figure 4 shows the most likely estimates of autocorrelation in the model, with the results summarized immediately below. Generally, we want to see the estimates clumped tightly to the left, near 0, with a mean estimate < .30. Such results suggest that the degree of autocorrelation is typical (according to research), and the adjustments necessary will be minor. Otherwise, it may be difficult to detect a change across phases.

#####Figure 4: Range of Possible Autocorrelation Coefficients

```{r, dpi=300, echo=FALSE}

#print(chains.JZS.AR[1:5,]) #Used to see the first five rows of mcmc output

hist(chains.JZS.AR[, 5], xlim = c(0, 1), xlab="Coefficient Estimate", main=NULL, col="slategrey")

summary(chains.JZS.AR[, 5])

```

### A. Is there a Credible DECREASE in the Level of Behavior from Phases A to B?

With an acceptable estimate of autocorrelation, we can safely test for changes in behavior that coincide with the phase change. If the intervention is intended to \*reduce\* a behavior following baseline (e.g., lower teacher ratings of problem behavior), we will apply a one-tailed, "left sided" test. The results of this analysis are summarized in Box 2.

##### Box 2. Bayes Factor Assessing Decrease in Behavior from Levels from Phase A to B

```{r, echo=FALSE}

output.JZS.AR=NULL

output.JZS.AR = ttest.Gibbs.AR(ypre, ypost, iterations = 20000, return.chains = TRUE, progress=FALSE, leftSided=TRUE, return.onesided=TRUE)

bf.leftsided = exp(output.JZS.AR$logbfOnesided)

print(bf.leftsided)

```

The first estimate ("rho acceptance rate") in Box 2 tells us about model efficiency. Estimates between .25 and .50 suggest that the model provides a good estimate of the intervention effect. If it is outside that range, increase the "iterations" value on line 119 of the markdown.

The next estimate in Box 2 is technically called the "JZS+AR Bayes Factor" (de Vries & Morey, 2013). Assuming the model is acceptable, the JZS+AR Bayes Factor tells us whether the data favor the null hypothesis (i.e., no differences between the phases in the anticipated direction) or the alternative hypothesis (i.e., the intervention phase coincided with changes from baseline in the anticipated direction).

Bayes Factors are interpreted in the following way:

BF Value | Interpretation

----------|------------------------------------------------------------------

10.00+ | Strong support for the null hypothesis

3.00+ | Moderate support for the null hypothesis

1.00 | No evidence in either direction (i.e., inconclusive)

< .30 | Moderate support for the alternative hypothesis (change occurred)

< .10 | Strong support for the alternative hypothesis (change occurred)

In the terminology of Bayesian statistics, estimates < .10 suggest that the change during intervention is "credible" (i.e., we can be very confident that change coincided with the intervention phase, and not just by chance).

Note that Bayes Factor estimates that exceed 1.0 suggests that the evidence supports the null hypothesis, with estimates exceeding 3.0 suggesting moderate support for the null, and estimates exceeding 10.0 suggesting strong support for the null (i.e., we can be very confident that change did NOT occur).

### B. Is there a Credible INCREASE in the Level of Behavior from Phases A to B?

If the intervention is intended to \*increase\* a behavior following baseline (e.g., improvements in student organization ratings), we apply a one-tailed, "right sided" test. The results of this analysis are in Box 4. The resulting rho acceptance rate and Bayes factor can be interpreted exactly as above, but the estimates are adjusted relative to the directional hypothesis.

##### Box 4. Bayes Factor Assessing Increase in Behavior from Levels from Phase A to B

```{r, echo=FALSE}

output.JZS.AR=NULL

output.JZS.AR = ttest.Gibbs.AR(ypre, ypost, iterations = 20000, return.chains = TRUE, progress=FALSE, leftSided=FALSE, return.onesided=TRUE)

bf.rightsided = exp(output.JZS.AR$logbfOnesided)

print(bf.rightsided)

```

### C. Estimating an Effect Size

The change from Phases A to B can be estimated as an effect size, but uncertaintly around that estimate creates a range of likely values. In Figure 5, the range of likely standardized effect sizes are ploted in a density plot. The 95% credible range is depicted as vertical blue lines. If the space between the blue lines does not include 0 on the x-axis, we can say with >95% confidence that the true effect is nonzero, and the sign of the value (negative or positve) tells us whether behavior decreased or increased in Phase B. Details regarding these estimates can be found below the figure.

One way to use this information is to continue interventions that have promising effect sizes until enough data are collected that the 95% credibility interval no longer includes zero. As more data are collected, the credible interval will narrow, assuming that the data continue to suggest the anticipated change.

#### Figure 5: Range of Likely Effect Sizes

```{r, echo=FALSE, dpi=300, warning=FALSE}

library(mcmcplots)

#print(chains.JZS.AR[1:5,]) #Used to see the first five rows of mcmc output

delta = chains.JZS.AR[,2]

denplot(delta, ci=0.95, main="", ylab="Density", xlab="Posterior Estimate")

summary(delta)

```

# III. Bayesian Single Case Analysis: Change in Intercepts and Slopes

If we want to see changes in behavioral trajectories from Phase A to Phase B, we can test for trend and intercept changes. This test will tell us if the intervention coincided with a meaningful change, either increasing or descreasing, in each phase. A change in intercept tells us about immediate change at phase shift (depending on the intent and nature of the intervention, these results may or may not be meaningful), and a change in the slope tells us about changes in trajectory from phases A to B. Figure 5 provides a visual representation of this test, with dashed lines representing trends in each phase.

##### Figure 5. Changes in Trend and Intercepts from Phase A to Phase B

```{r, dpi=300, echo=FALSE, warning=FALSE}

Trendplot <- ggplot(ssdall, aes(x=xrownum, y=myy, colour=myphase)) +

geom\_vline(linetype=2, xintercept=n1+.5)

Trendplot + geom\_line(size=1, na.rm=TRUE) + geom\_point(na.rm=TRUE) +

theme(legend.title=element\_blank()) +

geom\_smooth(data=ssdall, method=lm, se=FALSE, linetype="longdash", size=0.5, alpha=0.5, na.rm=TRUE) +

labs(x = "Time", y = "Data")

```

We begin by recalculating the autocorrelation in this new model.

```{r, echo=FALSE, results="hide"}

output.JZS.TAR = trendtest.Gibbs.AR(ypre, ypost, iterations = 10000, return.chains = TRUE, r.scaleInt = 1, r.scaleSlp = 1, betaTheta = 5, sdMet = 0.3, progress=FALSE)

chains.JZS.TAR = output.JZS.TAR$chains

```

#####Figure 5: Range of Possible Autocorrelation Coefficients

```{r, dpi=300, echo=FALSE}

#print(chains.JZS.TAR[1:5,]) #Used to see the first five rows of mcmc output

hist(chains.JZS.TAR[, 8], xlim = c(0, 1), xlab="Coefficient Estimate", main=NULL, col="slategrey")

summary(chains.JZS.TAR[, 8])

```

As before, the hope is that the autocorrelation estimate is generally below .30 to ensure that the estimates are not unduly penalized.

### A. Is there a Credible DECREASE in Intercepts/Slopes from Phases A to B?

With autocorrelation in the model accounted for, Bayes Factors can be estimated for the changes in intercepts and slopes. We start by testing whether there was a \*decrease\* in behavior coinciding with the phase change.

##### Box 5. Bayes Factors Assessing Decrease in Intercept and Slope from Phase A to B

```{r, echo=FALSE}

output.JZS.TAR=NULL

output.JZS.TAR = trendtest.Gibbs.AR(ypre, ypost, iterations = 10000, return.chains = TRUE, r.scaleInt = 1, r.scaleSlp = 1, betaTheta = 5, sdMet = 0.3, leftSidedInt=TRUE, leftSidedSlp=TRUE, return.onesided=TRUE, progress=FALSE)

bf.tarLeft = exp(output.JZS.TAR$logbfOnesided)

print(bf.tarLeft)

```

As before, the first estimate ("rho acceptance rate") in Box 5 tells us about model efficiency, with estimates between .25 and .50 suggesting that the model is reliable.

The next set of estimates in Box 5 are Bayes Factors, which tell us whether the evidence suggests that a nonrandom decrease in the slope and intercept coincided with the phase change, or if the evidence suggests that things stayed the same (or increased). We can interpret these estimates in the same way as before, with value > 1.0 suggesting that the evidence supports the null hypothesis (i.e., no change occurred in the anticipated direction), and values < .30 suggesting that the evidence suggests that credible changes occurred in Phase B.

### B. Is there a Credible INCREASE in Intercepts/Slopes from Phases A to B?

Alternatively, we can test for an increase in intercepts and slopes from phases A to B. Those results are summarized in Box 6.

##### Box 6. Bayes Factors Assessing Increase in Intercept and Slope from Phase A to B

```{r, echo=FALSE}

output.JZS.TAR=NULL

output.JZS.TAR = trendtest.Gibbs.AR(ypre, ypost, iterations = 10000, return.chains = TRUE, r.scaleInt = 1, r.scaleSlp = 1, betaTheta = 5, sdMet = 0.3, leftSidedInt=FALSE, leftSidedSlp=FALSE, return.onesided=TRUE, progress=FALSE)

bf.tarRight = exp(output.JZS.TAR$logbfOnesided)

print(bf.tarRight)

```

The estimates are interpreted similar to the examples above.

### C. Estimates of Effect Sizes

To make sense of these results, it may help to examine the estimates of change.

The density of likely differences from Phase A to Phase B slope is given in Figure 7 below. The density plot shows the most likely estimates of how much change occurred in the slope. The "confidence interval" in Bayesian statistics is referred to as the highest density interval (HDI), and it is denoted by the two vertical blue lines. If the HDI includes 0 in Figure 5, we cannot say with confidence that a change in slope occurred coinciding with the shift between phases. If the HDI does NOT cross zero, we can be > 95% confident that a nonrandom, nonzero slope change coincided with the switch to Phase B.

#####Figure 7: Posterior Distribution of Likely Changes in Slope

```{r, dpi=300, echo=FALSE, message=FALSE, warning=FALSE}

denplot(chains.JZS.TAR[, 4], ci=0.95, main="", ylab="Density", xlab="Posterior Estimate")

summary(chains.JZS.TAR[, 4])

```

(Note that the estimate above in the unstandardized metric of the outcome measure, and it may vary somewhat from the regression lines in Figure 5 due to the adjustment for autocorrelation.)

Similarly, we can test the change in intercepts from one phase to the next. Generally speaking, changes in intercepts are not important if there are clear changes in trends. In cases where there are no trend differences, a change in the intercept would suggest that a jump in performance coincided with the phase change, either up or down.

#####Figure 8: Posterior Distribution of Likely Changes in Intercepts

```{r, dpi=300, echo=FALSE, message=FALSE, warning=FALSE}

denplot(chains.JZS.TAR[, 2], ci=0.95, main="", ylab="Density", xlab="Posterior Estimate")

summary(chains.JZS.TAR[, 2])

```

(Note that the estimate above in the unstandardized metric of the y-axis, and it may vary somewhat from the regression lines in Figure 5 due to the adjustment for autocorrelation.)

As before, these effect size estimates might be used to decide how long to continue a promising intervention. Once the 95% credibility intervals do not cross zero, we can safely conclude that a nonrandom change occurred, coinciding with the shift to the intervention phase.

### References

Auerbach, C., & Zeitlin, W. (2014). \*SSD for R: An R Package for Analyzing Single-Subject Data\*. New York: Oxford.

de Vries, R.M., Hartogs, B.M., & Morey, R.D. (2015). A tutorial on computing Bayes factors for single-subject designs. \*Behavior Therapy\*, \*46\*, 809-823.

de Vries, R.M., & Morey, R.D. (2013). Bayesian hypothesis testing for single-subject designs. \*Psychological Methods\*, \*18\*, 165-185. doi: 10.1037/a0031037

Riley-Tillman, T.C., & Burns, M.K. (2009). \*Evaluating Educational Interventions: Single-Case Design for Measuring Response to Intervention.\* New York: Guilford Press.