# BDA - Assignment 8

*Anonymous*

## Model assessment: LOO-CV for factory data with Stan

In this assignment, we are going to apply **leave-one-out cross-validation (LOO-CV)** to assess the predictive performance of the pooled, separate and hierarchical Gaussian models for the factory dataset. Let us run the stan simulations from the last week's assignment.

### 1. Run the stan-models

```
data("factory")
```

### i. The separate model

```
N <- ncol(factory) * nrow(factory)
K <- ncol(factory)
x <- rep(1:ncol(factory), nrow(factory))
y <- factory

df <- list(N=N, K=K, y=c(t(y)), x=x)
fit_separate <- stan(file="mod_sep.stan", data=df, verbose = FALSE)
#separate_model <- extract(fit_separate)
```

### ii. The pool model

```
fit_pooled <- stan(file="mod_pooled.stan", data=df, verbose = FALSE)
```

```
## Warning in readLines(file, warn = TRUE): incomplete final line found on 'C:
## \Users\alisa_000\Documents\BDA\mod_pooled.stan'
```

```
#pooled_model <- extract(fit_pooled)
```

### iii. The hierarchial model

```
fit_hierarchical <- stan(file='mod_hier.stan', data=df)
```

```
## Warning in readLines(file, warn = TRUE): incomplete final line found on 'C:
## \Users\alisa_000\Documents\BDA\mod_hier.stan'
```

```
## Warning: There were 278 divergent transitions after warmup. Increasing adapt_delta above 0.8 may help
## http://mc-stan.org/misc/warnings.html#divergent-transitions-after-warmup
```

```
## Warning: Examine the pairs() plot to diagnose sampling problems
```

```
## Warning: The largest R-hat is 1.19, indicating chains have not mixed.
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#r-hat

## Warning: Bulk Effective Samples Size (ESS) is too low, indicating posterior means and medians may be
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#bulk-ess

## Warning: Tail Effective Samples Size (ESS) is too low, indicating posterior variances and tail quant
## Running the chains for more iterations may help. See
## http://mc-stan.org/misc/warnings.html#tail-ess
```

```
#h <- extract(fit_hierarchical)
```

Please do not mind the warnings!

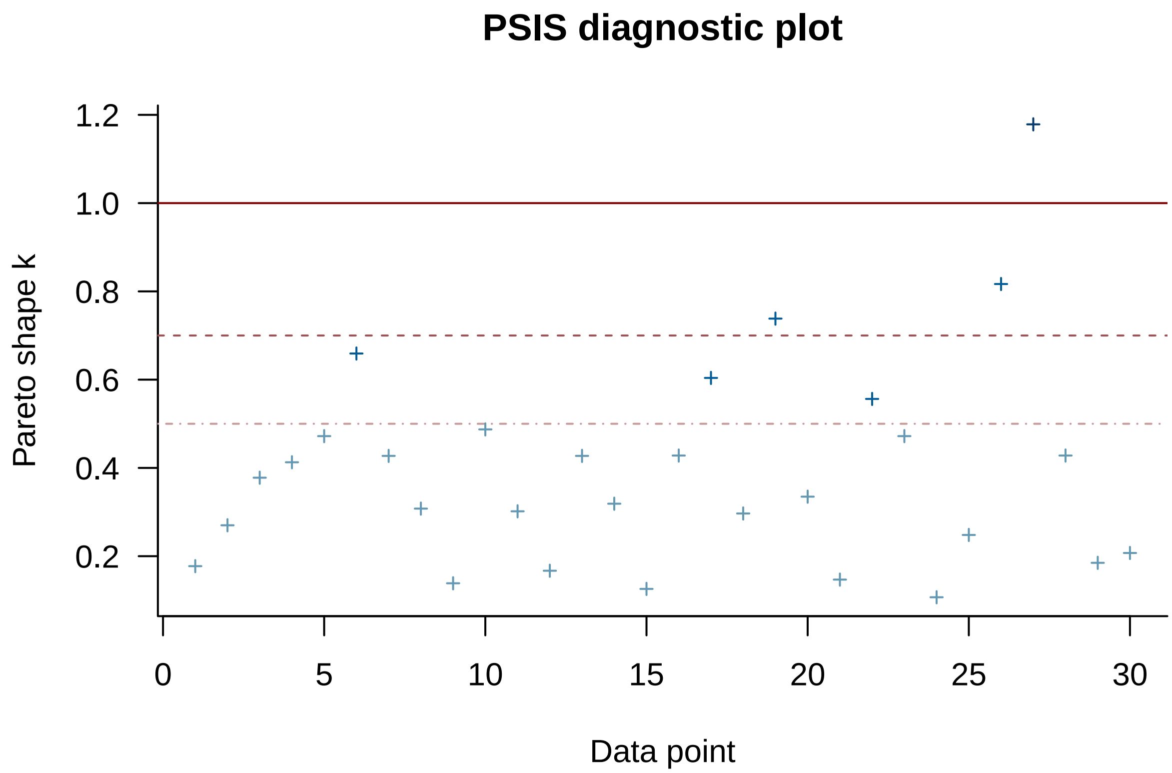
**2. Compute the PSIS-LOO elpd values and the k-values for each of the three models**

```
# The log-likelihood array matrix and vector of effecive sample size estimates needed for each model

ll_sep <- extract_log_lik(fit_separate, merge_chains = F)
eff_sep <- relative_eff(exp(ll_sep))
loo_sep <- suppressWarnings(loo(ll_sep, eff_sep))
loo_sep
```

```
##
## Computed from 4000 by 30 log-likelihood matrix
##
##          Estimate   SE
## elpd_loo   -130.8  6.1
## p_loo        13.3  3.3
## looic       261.6 12.2
## ------
## Monte Carlo SE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     24    80.0%   1488
##  (0.5, 0.7]   (ok)        3    10.0%   207
##    (0.7, 1]   (bad)       2     6.7%   55
##    (1, Inf)   (very bad)  1     3.3%   8
## See help('pareto-k-diagnostic') for details.
```

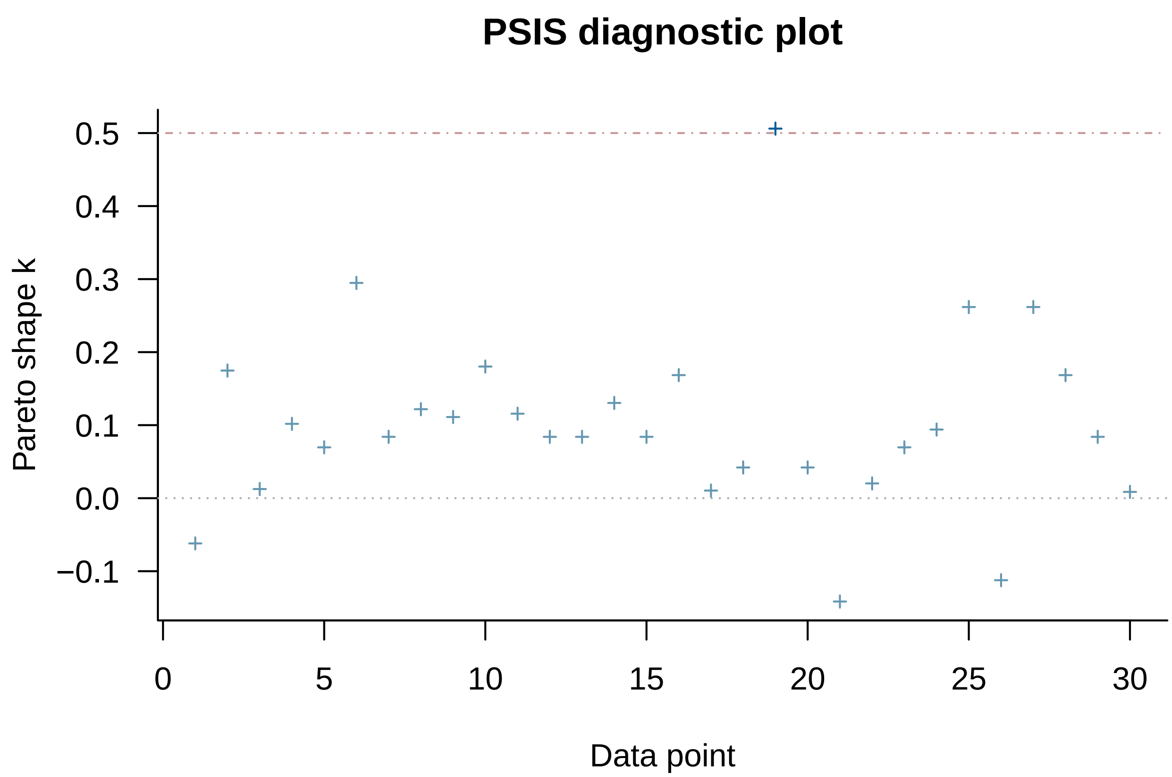
```
# Diagnostics plot
plot(loo_sep)
```




```r
ll_pool <- extract_log_lik(fit_pooled, merge_chains = F)
eff_pool <- relative_eff(exp(ll_pool))
loo_pool <- suppressWarnings(loo(ll_pool, eff_pool))
loo_pool
```

```
##
## Computed from 4000 by 30 log-likelihood matrix
##
##          Estimate  SE
## elpd_loo   -131.1 4.7
## p_loo         2.3 1.0
## looic       262.2 9.4
## ------
## Monte Carlo SE of elpd_loo is 0.0.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)     29    96.7%   2419
##  (0.5, 0.7]   (ok)        1     3.3%   800
##    (0.7, 1]   (bad)       0     0.0%   <NA>
##    (1, Inf)   (very bad)  0     0.0%   <NA>
##
## All Pareto k estimates are ok (k < 0.7).
## See help('pareto-k-diagnostic') for details.
```

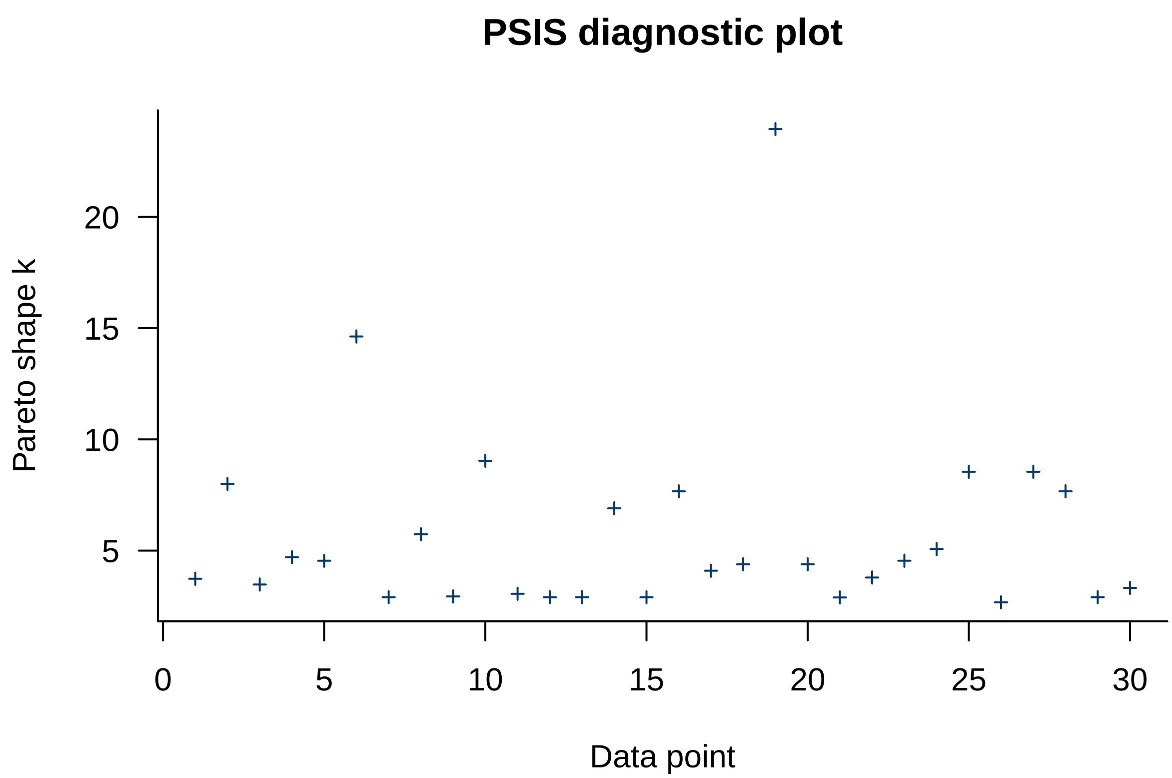
```r
plot(loo_pool)
```




```r
ll_hier <- extract_log_lik(fit_hierarchical, merge_chains = F)
eff_hier <- relative_eff(exp(ll_hier))
loo_hier <- suppressWarnings(loo(ll_hier, eff_hier))
loo_hier
```

```
##
## Computed from 4000 by 30 log-likelihood matrix
##
##          Estimate    SE
## elpd_loo  -5575.6 461.2
## p_loo      1963.1 367.6
## looic     11151.2 922.4
## ------
## Monte Carlo SE of elpd_loo is NA.
##
## Pareto k diagnostic values:
##                          Count Pct.    Min. n_eff
## (-Inf, 0.5]   (good)      0      0.0%   <NA>
##  (0.5, 0.7]   (ok)        0      0.0%   <NA>
##    (0.7, 1]   (bad)       0      0.0%   <NA>
##    (1, Inf)   (very bad) 30    100.0%   1
## See help('pareto-k-diagnostic') for details.
```

```r
plot(loo_hier)
```




My first plot looked normal (the range was from 0 to 1), but then suddenly I started to get these weird plot. My answers below are based on the first plot I obtained as I believe it was the correct one.

### 3. Compute the effective number of parameters peff for each of the three models

More on theory refer to BDA page 176. These estimates (the effective number of parameters) can be directly obtained from the output of the function above. The output also shows the standard error related to the pointestimate.

```r
# For separate model
loo_sep$estimates[2,1]
```

```
## [1] 13.28987
```

```r
# For pooled model
loo_pool$estimates[2,1]
```

```
## [1] 2.314601
```

```r
# For hier model
loo_hier$estimates[2,1]
```

```
## [1] 1963.053
```

#### 4. Assess how reliable the PSIS-LOO estimates are for the three models based on the k-values

In terms of the PSIS-LOO analysis, separate model did not perform well. As there were multiple k-values over 0.7, we might be overestimating the predictive accuracy of the model. Whereas pooled model and hierarchial model had all of their k-values below the 0.7 threshold, we could state that PSIS-LOO estimate of these models (hierarchial and pooled) can be considered to be reliable.

#### 5. Which model should be selected according to PSIS-LOO?

Definitely the hierarchial model as the elpd is the highest for it which means the goodness-of-fit is on pretty good level. Also the PSIS-LOO estimates are quite reasonable for hierarchial model.