bootstrap_fragment

August 8, 2024

```r
# Load necessary library
library(MASS)

# Read data from file
data <- read.table("pep1_ws17-pep2_ss18-nozeros.txt", header = TRUE)

# Calculate the original correlation coefficient and its confidence interval
original_cor <- cor.test(data[,1], data[,2])
original_ci <- c(original_cor$conf.int[1], original_cor$conf.int[2])

# Total number of data points (students)
n <- nrow(data)

# Define the number of bootstrap samples to use
# (using the suggestion in the lecture notes)
# (but you can try different numbers to see the effects)
n_bootstrap <- ceiling(n * log(n)^2)

# Initialize vector to store bootstrap correlations
bootstrap_correlations <- numeric(n_bootstrap)
```

```r
# Your bootstrap sampling routine goes here
for (i in 1:n_bootstrap) {
  # Resample data with replacement
  #.............code.............
  # Calculate correlation for resampled data
  #.............code.............
}
```

```r
# Calculate bootstrap confidence intervals
ci_lower <- quantile(bootstrap_correlations, 0.025)
ci_upper <- quantile(bootstrap_correlations, 0.975)
```

```r
# Print results
cat("Original correlation coefficient        : ", round(original_cor$estimate,
  3), "\n")
cat("95% confidence interval original sample : ", round(original_ci, 3), "\n")
```

```r
cat("Correlation coefficient bootstrap sample: ",
 ↪round(mean(bootstrap_correlations), 3), "\n")
cat("95% confidence interval bootstrap sample: ", round(ci_lower, 3),
 ↪round(ci_upper, 3), "\n")
```

```r
# Plot histogram of bootstrap correlations
hist(bootstrap_correlations, breaks = 50, main = "Bootstrap Correlations", xlab
 ↪= "Correlation Coefficients")
abline(v = original_cor$estimate, col = "red", lwd = 2, lty = 2)
abline(v = ci_lower, col = "blue", lwd = 2, lty = 2)
abline(v = ci_upper, col = "blue", lwd = 2, lty = 2)
```

[ ]: