



NAME: _____

1 Objective: Studentized and Wild Bootstrapping

Our goal is the same as in Lab 19, producing graphs like those of Mammen, distributed in class, when sampling from two populations. However, this time we focus on studentized and wild bootstrapping. NOTE: Recall that a function $g(X, \theta)$ of data and parameters is said to be a "pivot" or a "pivotal quantity" if its distribution does not depend on the parameter. One primary example of a pivotal quantity

$$\text{is } g(X, \mu) = \frac{\bar{X}_n - \mu}{S_n / \sqrt{n}}$$

2 Key Words

rchisq, quintile, sample, t, apply, cbind, pnorm, studentized bootstrap, wild bootstrap.

3 Stu, Wild, and Stu-Wild Bootstraps

Basic definitions and the notations are the same as in Lab 19 (see Brief Description).

[A] STU BOOTSTRAP: Studentized bootstrap is concerned with the studentized pivot $Q_s = (\hat{M} - M) / \hat{D}$ and its bootstrap version Q_s^* .

Here M denotes population mean, \hat{M} sample mean and \hat{D} sample standard deviation. we are still after BIAS and MSE, but now there is a slight change: studentization.

- [1] Modify the inner loop of `bs.p.***` to give studentized bootstrap estimates of BIAS & MSE: Function `bs.p.***` becomes `bs.s.***`. Below is the part that needs modification:

```
# ... outer loop as before
for(j in 1:nr) {
  ss <- sample(s, ns, rep = T)      # get S*
  sd <- sqrt(var(ss)/ns)           # get sd*
  dm <- (mean(ss) - ms)/sd         # stu diff in mean
  fr <- fr + (dm < uu)             # freq
}
# ... as before
```

- [2] Use the function to get plot data

```
bss8 <- bs.s.chi8(20, 50, 100, uks.8)           # NOTE: uks.8 NOT uk.8
```

[3] Make plots as in Lab 19 [B3].

[B] Start with a population P , select a sample S , size ns from P and form the pivot $Q = \widehat{M} - M :=$ sample mean – population mean.

KEY IDEA: The distribution of Q , **across various samples from P** , should be close to the distribution of $Q^* = Q^* | S$, **across various samples from S** .

PLAIN BOOTSTRAP chooses nr samples of size ns , with replacement, from S (say S_1^*, \dots, S_{nr}^*) and for each calculates the plug-in pivot Q^* (say Q_1^*, \dots, Q_{nr}^*). Specifically, if $ns = 10$ and if $S = \{x_1, \dots, x_{10}\}$, one sample from S might be S_1^* which contains one x_2 , four x_5 , and five x_9 . Then

$$\begin{aligned} Q_1^* = M^* - \widehat{M} &= (x_2 + 4x_5 + 5x_9)/10 - \widehat{M} \\ &= \left[1(x_2 - \widehat{M}) + 4(x_5 - \widehat{M}) + 5(x_9 - \widehat{M}) \right] / 10 \end{aligned}$$

WILD BOOTSTRAP: Just resamples from S differently, by choosing a_1 x_1 's, \dots , a_{10} x_{10} 's, where a_1, \dots, a_{10} are independent Poisson random variables: Now the pivot is

$$\begin{aligned} Q_1^* &= \left[a_1(x_1 - \widehat{M}) + \dots + a_{10}(x_{10} - \widehat{M}) \right] / 10 \\ &= \left[(a_1 - 1)(x_1 - \widehat{M}) + \dots + (a_{10} - 1)(x_{10} - \widehat{M}) \right] / 10 \text{ (Trick1: not used, } \lambda = 1) \\ &= \left[(a_1 - \text{mean}(a))x_1 + \dots + (a_{10} - \text{mean}(a))x_{10} \right] / 10 \text{ (Trick2)} \end{aligned}$$

[1] Modify the inner loop of `bs.p.***` to give wild bootstrap estimates of BIAS & MSE:

Function `bs.p.***` becomes `bs.w.***`. Below is the part that needs modification:

```
# ... outer loop as before
for(j in 1:nr) {
  ak <- rpois(ns, 1)
  cak <- ak - mean(ak)           # last line in above equation, trick 2
  dm <- mean(cak * s)           # Trick 2: MW
  fr <- fr + (dm < uu)          # frequency
}
# ... as before
```

[2] Use the function to get plot data

```
bsw8 <- bs.w.chi8(20,50,100,uk.8)           # NOTE uk.8
```

[3] Make plots as in Lab 19 [B3].

[C] WILD STU BOOTSTRAP: Just like plain studentized, but formulas worse (see Hall and Mammen, Annals of Statistics, 1994, p.2017).

[1] Modify the inner loop of `bs.p.***` to give wild studentized bootstrap estimates of BIAS & MSE:

Function `bs.p.***` becomes `bs.ws.***`. Below is the part that needs modification:

```
# ... storage as before
for(i in 1:nb) {
  s <- . . .
  ms <- mean(s)
  ms2 <- mean((s - ms)^2)
  fr <- rep(0, 19)

  for(j in 1:nr) {
    ak <- rpois(ns, 1)
    cak <- ak - mean(ak)
    w1 <- mean(cak * s)
    w2 <- mean(cak * s^2)
    sig2 <- ms2 + w2 - w1^2 - 2*ms*w1
    dm <- w1 / sqrt(sig2)
    fr <- fr + (dm < uu)
  }
# ... as before
}
```

[2] Use the function to get plot data

```
bsws <- bs.ws.chi8(20, 50, 100, uks.8)      # NOTE: uks.8 NOT uk.8
```

[3] Make plots and compare.