STAT 460/560: Statistical Inference 1 **Winter Term 1 (Fall) 2026**

Assignment 4

Due date: Oct. 9, 11 p.m. *Instructor: Daniel McDonald and Benjamin Bloem-Reddy*

Instructions

Academic integrity policy: I encourage you to discuss verbally with other students about the assignment. However, you **must** write your answers by yourself. If you discuss the assignment with anyone (a classmate or anyone else), you must say so at the top of your solutions. Use of AI tools to produce full or partial answers to these questions or to related questions is strictly prohibited. This includes tools like Co-pilot, auto-complete, etc. Also, refrain from looking at answer keys from other schools, previous years, or the Math Stack Exchange. For more information, see: the syllabus or contact the instructor.

1. Solutions may be handwritten or typeset with LaTeX (see the course website for a template). It must be legible. Difficult-to-read solutions will be marked down or ignored.
2. *Justify your answers formally. (Err on the side of too much detail so that the teaching team knows exactly what you're doing and why.)*
3. Submit your work by 11:59 pm (Vancouver time) on the due date to Canvas (.pdf).
4. For questions that require computer simulations, you may use any language from the list of `(R, Julia, Python)`, with level of support provided by the teaching team decreasing further in the list. Please include in your PDF submission the relevant plots generated by your simulation code, and append your code (in PDF) to your submission. Please retain the code you used, as you may be asked for it if required.

Questions

1. ~~Wasserman (2006) 2.8. (Same as Wasserman 2004, 7.6 on assignment 3.)~~
2. Wasserman (2006) 2.10. (Note that this should state that $\theta = T(F) = F^{-1}(p)$ is the p-th quantile of F.)

   Solution:

   Note, this is a bit more finicky than intended. I've tried to be as careful as possible, but I don't expect student solutions to be this careful.

   Assume F is differentiable at θ with $f(\theta) = F'(\theta) > 0$, and that F is strictly increasing near θ, so F^{-1} is continuous at p and $(F^{-1})'(p) = 1/f(\theta)$. Define $T(G) = \inf\{y : G(y) \geq p\}$. Let $F_\epsilon = (1-\epsilon)F + \epsilon\,\delta_x$ for $\epsilon \in [0, 1)$, where $\delta_x(y) = \mathbf{1}\{y \geq x\}$, and write $T_\epsilon = T(F_\epsilon)$. Since $\epsilon \geq 0$, the influence function is the right derivative $L(x) = \frac{d}{d\epsilon} T_\epsilon \big|_{\epsilon=0^+}$. (The differentiation is done on a smooth equation in each case below, because $\delta_x(T_\epsilon)$ is a step function of T_ϵ and cannot be differentiated by the chain rule.)

Case $x < \theta$. Since F^{-1} is continuous at p, $T_\epsilon \to \theta > x$, so for small ϵ we have $T_\epsilon > x$ and $\delta_x(T_\epsilon) = 1$. The defining equation becomes

$$p = (1-\epsilon)F(T_\epsilon) + \epsilon \quad \Longrightarrow \quad F(T_\epsilon) = \frac{p-\epsilon}{1-\epsilon} \quad \Longrightarrow \quad T_\epsilon = F^{-1}\Big(\frac{p-\epsilon}{1-\epsilon}\Big).$$

Since $\frac{d}{d\epsilon}\frac{p-\epsilon}{1-\epsilon}\Big|_{\epsilon=0} = p-1$, the chain rule gives

$$L(x) = \frac{p-1}{f(\theta)}.$$

Case $x > \theta$. Similarly $T_\epsilon < x$ for small ϵ, so $\delta_x(T_\epsilon) = 0$ and

$$p = (1-\epsilon)F(T_\epsilon) \quad \Longrightarrow \quad T_\epsilon = F^{-1}\Big(\frac{p}{1-\epsilon}\Big), \qquad L(x) = \frac{p}{f(\theta)}.$$

Case $x = \theta$. *(extra care needed, but potentially omitted from student solutions.)*
Here the argument above fails, since T_ϵ may lie on either side of x. We compute T_ϵ directly. For $y < \theta$ we have $F(y) \le p$, so $F_\epsilon(y) = (1-\epsilon)F(y) \le (1-\epsilon)p < p$. At $y = \theta$,

$$F_\epsilon(\theta) = (1-\epsilon)p + \epsilon = p + \epsilon(1-p) \ge p.$$

Hence $T_\epsilon = \theta$ for every $\epsilon \in (0,1)$: the atom at θ absorbs the perturbation and the quantile does not move. (Note that $F_\epsilon(T_\epsilon) = p$ fails here, because F_ϵ jumps at θ; this is why the equation $p = F_\epsilon(T_\epsilon)$ cannot be used.) Therefore

$$L(\theta) = 0.$$

Conclusion.

$$L(x) = \begin{cases} \dfrac{p-1}{f(\theta)}, & x < \theta, \\ 0, & x = \theta, \\ \dfrac{p}{f(\theta)}, & x > \theta. \end{cases}$$

Wasserman's statement, $L(x) = \frac{p-1}{f(\theta)}$ for $x \le \theta$ and $\frac{p}{f(\theta)}$ for $x > \theta$, agrees with this except at the single point $x = \theta$. The value there does not affect anything that matters for inference, since $\{\theta\}$ has F-measure zero. In particular,

$$\int L\, dF = \frac{(p-1)p}{f(\theta)} + \frac{p(1-p)}{f(\theta)} = 0$$

and $\int L^2\, dF = \frac{p(1-p)}{f(\theta)^2}$, so the asymptotic variance of the sample quantile is unchanged.

3. Wasserman (2006) 3.9. Assume that $X_1, \ldots, X_n$ are i.i.d. with mean μ and variance σ^2.

Solution:

Since the bootstrap sample is drawn with replacement from the original sample, we have:

$$\begin{aligned} \mathbb{E}(\overline{X}^* \mid X_1, \ldots, X_n) &= \mathbb{E}\left(n^{-1} \sum_{i=1}^n X_i^* \;\middle|\; X_1, \ldots, X_n \right) \\ &= n^{-1} \sum_{i=1}^n \mathbb{E}(X_i^* | X_1, \ldots, X_n) \\ &= n^{-1} \sum_{i=1}^n X_i \\ &= \bar{X}. \end{aligned}$$

Note that the second-to-last equality follows from the fact that each X_i^* has equal probability of being any of the X_i's, so the conditional expectation is just the sample average of the X_i's.

Similarly,

$$\begin{aligned}
\text{Var}(\overline{X}^* \mid X_1, \ldots, X_n) &= \text{Var}\left(n^{-1} \sum_{i=1}^n X_i^* \;\middle|\; X_1, \ldots, X_n \right) \\
&= n^{-2} \sum_{i=1}^n \text{Var}(X_i^* | X_1, \ldots, X_n) \\
&= n^{-2} \sum_{i=1}^n \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X})^2 \\
&= \frac{n-1}{n} \frac{s^2}{n}
\end{aligned}$$

where s^2 is the sample variance of X_i.

For the unconditional moments,

$$\mathbb{E}(\overline{X}^*) = \mathbb{E}[\mathbb{E}(\overline{X}^* \mid X_1, ..., X_n)] = \mathbb{E}(\overline{X}) = \mu,$$

and

$$\begin{aligned}
\text{Var}(\overline{X}^*) &= \mathbb{E}[\text{Var}(\overline{X}^* \mid X_1, ..., X_n)] + \text{Var}(\mathbb{E}[\overline{X}^* \mid X_1, ..., X_n]) \\
&= \frac{n-1}{n} \frac{\mathbb{E}[s^2]}{n} + \text{Var}(\overline{X}) \\
&= \frac{n-1}{n} \frac{\sigma^2}{n} + \frac{1}{n} \sigma^2 \\
&= \frac{2n-1}{n} \frac{\sigma^2}{n}.
\end{aligned}$$

4. Wasserman (2006) 3.9. By "parametric bootstrap", we mean that you should assume that the data are distributed according to the model, but with μ unknown. So given the data, you should estimate μ by $\hat{\mu} = \bar{X}$, and then generate samples from the distribution $N(\hat{\mu}, 1)$. This is in contrast to the "nonparametric bootstrap" where you would just sample with replacement from the original data.

   Use $B = 10000$ bootstrap samples. And use the Pivotal interval the bootstrap.

   Solution:

   By the delta method, we have

$$\text{Var}(\hat{\theta}) \approx \left(\frac{d}{d\mu} e^{\mu} \right)^2 \text{Var}(\bar{X}) = e^{2\mu} \frac{1}{n}.$$

   Thus, the standard error of $\hat{\theta}$ is approximately

$$\text{SE}(\hat{\theta}) \approx \frac{e^{\mu}}{\sqrt{n}}.$$

   A 95% confidence interval for θ is then

$$\hat{\theta} \pm 1.96 \cdot \text{SE}(\hat{\theta}) \approx e^{\bar{X}} \pm 1.96 \cdot \frac{e^{\bar{X}}}{\sqrt{n}} = e^{\bar{X}} \left(1 \pm \frac{1.96}{\sqrt{n}} \right).$$

   Note that $\overline{X} \sim N(\mu, 1/n)$. Therefore $\exp(\overline{X})$ is log-normally distributed.

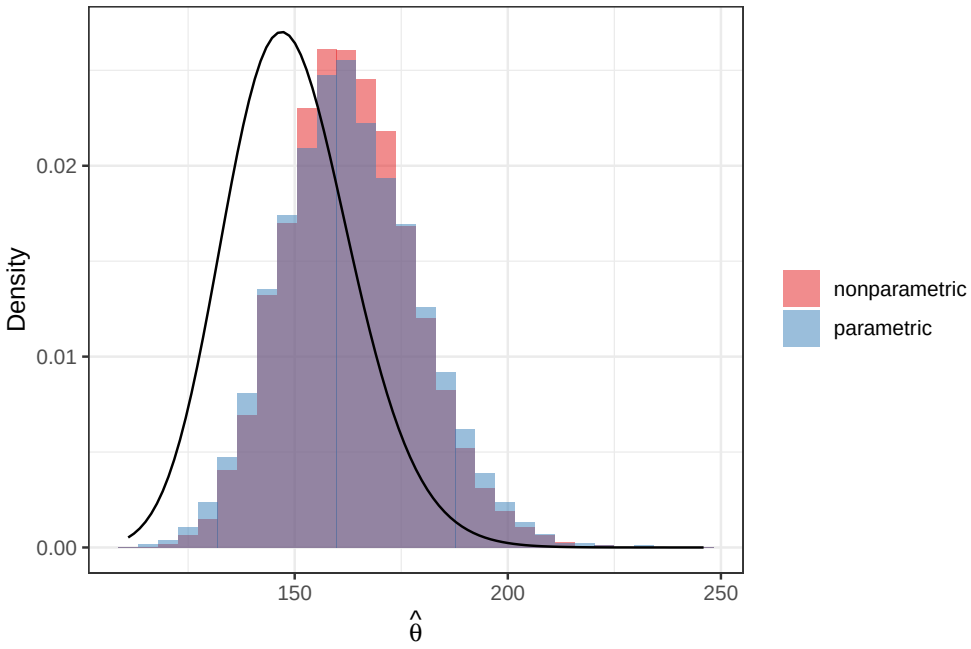


Figure 1: Result for Exercise 3.10: Parametric and nonparametric bootstrap distributions of $\hat{\theta} = \exp(\bar{X})$ with $n = 100$ and $\mu = 5$. The black curve is the density of the log-normal distribution of $\hat{\theta}$. Not that it is "off" a bit due to the fact that $\exp(\bar{X})$ is not that similar to $\exp(\mu)$ for this seed.

```r
suppressMessages(library(tidyverse))

% ## Create our data
mu <- 5
n <- 100
B <- 10000
alpha <- 0.05
set.seed(123)
X <- rnorm(n, mean = mu, sd = 1)
```

```r
# (a)
xbar <- mean(X)
theta_hat <- exp(xbar)

delta_method <- c(
  theta_hat *
    (1 + c(-1, 1) * (qnorm(1 - alpha / 2) / sqrt(n))) |>
      set_names(stats:::format_perc(c(alpha / 2, 1 - alpha / 2))),
  se = theta_hat / sqrt(n)
)

# Parametric bootstrap
para_boot <- map_dbl(
  1:B,
```

```
  ~ {
    X_boot <- rnorm(n, mean = xbar, sd = 1)
    exp(mean(X_boot))
  }
)
para_boot_stats <- c(
  2 *
    theta_hat -
    quantile(para_boot, probs = c(1 - alpha / 2, alpha / 2)) |>
      set_names(stats:::format_perc(c(alpha / 2, 1 - alpha / 2))),
  se = sqrt((n - 1) / n) * sd(para_boot)
)

# Nonparametric bootstrap
np_boot <- map_dbl(
  1:B,
  ~ {
    X_boot <- sample(X, size = n, replace = TRUE)
    exp(mean(X_boot))
  }
)
np_boot_stats <- c(
  2 *
    theta_hat -
    quantile(np_boot, probs = c(1 - alpha / 2, alpha / 2)) |>
      set_names(stats:::format_perc(c(alpha / 2, 1 - alpha / 2))),
  se = sqrt((n - 1) / n) * sd(np_boot)
)
bind_rows(
  delta_method = delta_method,
  para_boot = para_boot_stats,
  np_boot = np_boot_stats,
  .id = "method"
)
```

```
# A tibble: 3 x 4
  method       `2.5%` `97.5%`    se
  <chr>         <dbl>   <dbl> <dbl>
1 delta_method   131.    194.  16.2
2 para_boot      128.    192.  16.2
3 np_boot        130.    189.  15.0
```

```
# (b)
histo <- bind_cols(parametric = para_boot, nonparametric = np_boot) |>
  pivot_longer(everything(), names_to = "method", values_to = "theta") |>
  ggplot(aes(x = theta)) +
  geom_histogram(
    aes(y = after_stat(density), fill = method),
    bins = 30,
    alpha = 0.5,
```

```
    position = "identity"
  ) +
  labs(x = expression(hat(theta)), y = "Density", fill = "") +
  scale_fill_brewer(palette = "Set1") +
  stat_function(
    fun = dlnorm,
    args = list(meanlog = mu, sdlog = 1 / sqrt(n)),
    color = "black"
  ) +
  theme_bw()

ggsave(
  histo,
  file = here::here("assignments", "img", "wasserman2006-3p10-histo.pdf"),
  width = 6,
  height = 4
)
```

See Figure 1 for the histogram of the parametric and nonparametric bootstrap distributions.

5. Wasserman (2006) 3.11(b). Feel free to use (a) as practice, but you do not need to submit it.

References

Larry Wasserman. *All of Statistics*. Springer, New York, 2004.

Larry Wasserman. *All of Nonparametric Statistics*. Springer, New York, 2006.