Statistics 583, Problem Set 9
Wellner; 5/25/2016

Reading: Chapter 8, sections 8.1- 8.4;
van der Vaart, Asymptotic Statistics, chapter 23, pages 326 - 340;
Wasserman, Chapters 2-3, pages 13-41.

Due: Wednesday, June 1, 2016

1. Suppose that $T(F) = Var_F(X)$ so that $T_n \equiv T(\mathbb{F}_n) = n^{-1}\sum_{i=1}^n (X_i - \overline{X})^2$. Show that the jackknife estimate of the variance $\sigma_n^2(F) \equiv Var_F(T_n)$ is

$$\widehat{Var} = \frac{n^2}{(n-1)^3}(\widehat{\mu}_4 - \widehat{\mu}_2^2)$$

where $\widehat{\mu}_k \equiv n^{-1}\sum_{i=1}^n (X_i - \overline{X})^k$ for $k = 1, 2, \ldots$. Hence, assuming that $EX^4 < \infty$, the jackknife estimate of variance is consistent for this T:

$$n\widehat{Var} \rightarrow_p \mu_4 - \mu_2^2 = \mu_2^2\{2 + \frac{\mu_4}{\mu_2^2} - 3\} = T_2(F)(2 + \gamma_2)\,.$$

2. (a) Wasserman, problem 3.8.9, page 40: Let $X_1, \ldots, X_n$ be n distinct observations (no ties). Let $X_1^*, \ldots, X_n^*$ denote a bootstrap sample (from the empirical d.f. $\mathbb{F}_n$ of the X_i's), and let $\overline{X}_n^* = n^{-1}\sum_{i=1}^n X_i^*$. Find: $E\{\overline{X}_n^*|X_1, \ldots, X_n\}$, $Var(\overline{X}_n^*|X_1, \ldots, X_n)$, and $Var(\overline{X}_n^*)$.
(b) Wasserman, problem 3.8.13, page 41: Let $X_1, \ldots, X_n$ be n distinct observations (no ties). Let $X_1^*, \ldots, X_n^*$ denote a bootstrap sample (from the empirical d.f. $\mathbb{F}_n$ of the X_i's). Let G denote the marginal distribution of X_i^*. Note that $G(x) = P(X_i^* \le x) = E\{P(X_i^* \le x|X_1, \ldots, X_n) = E\{\mathbb{F}_n(x)\} = F(x)$. So it appears that X_i^* and X_i have the same distribution. But in (a) we showed that $Var(\overline{X}_n) \ne Var(\overline{X}_n^*)$. Explain.

3. Consider a $V-$functional of order $r = 2$ given by $T(P) = \int\int h(x, y)dP(x)dP(y)$ where h is permutation symmetric. Find the jackknife estimate of bias for the ($V-$statistic) estimator $T(\mathbb{P}_n)$ of $T(P)$. Also find the jackknife estimator of $T(P)$.

4. Wasserman, problem 3.8.11, page 41: Let $X_1, \ldots, X_n$ be i.i.d. Uniform$(0, \theta)$. The MLE of θ is $\widehat{\theta}_n \equiv X_{(n)} = \max\{X_1, \ldots, X_n\}$.
(a) Find the distribution of $\widehat{\theta}_n$ and the exact and limiting distribution of $n(\theta - \widehat{\theta}_n)$.
(b) Compare the true and limiting distribution of $n(\theta - \widehat{\theta}_n)$ with the parametric and nonparametric bootstrap distributions when $\theta = 1$.
(c) Show that for the parametric bootstrap $P(\hat{\theta}_n^* = \hat{\theta}_n) = 0$ but for the nonparametric bootstrap $P(\hat{\theta}_n^* = \hat{\theta}_n) = 1 - (1 - 1/n)^n \rightarrow 1 - e^{-1} \approx .632\ldots$.

5. **Optional bonus problem 1:** (Bootstrapping a linear regression model a simple way.) Consider bootstrapping a linear regression model

$$Y_i = \mathbf{x}_i^T \beta + \epsilon_i, \qquad i = 1, \ldots, n$$

where the ϵ_i are i.i.d. mean 0, finite variance, and the $\mathbf{x}_i$ are given $p-$dimensional vectors, such that there is no constant term in the regression.
(a) Show that the estimated residuals $\hat{\epsilon}^T = (\hat{\epsilon}_1, \ldots, \hat{\epsilon}_n)$ satisfy $\hat{\epsilon} - \epsilon = -H\epsilon$ where $H = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ is the "hat matrix" (i.e. the projection matrix onto the column space of $\mathbf{X}$).
(b) Suppose that $\hat{\epsilon}_1^*, \ldots, \hat{\epsilon}_n^*$ is a bootstrap sample (with replacement) from $\{\hat{\epsilon}_1, \ldots, \hat{\epsilon}_n\}$. Show that

$$E_*(n^{1/2}(\hat{\beta}^* - \hat{\beta})) = (\frac{1}{n}\mathbf{X}^T\mathbf{X})^{-1}(\frac{1}{n}\mathbf{X}^T\mathbf{1})Z_n$$

where $Z_n = n^{-1/2}\sum_{i=1}^n \hat{\epsilon}_i$.
(c) Show that if $\max_{1\le i\le n} h_{ii} \to 0$, and $n^{-1}\mathbf{X}^T\mathbf{X} \to V$, a positive definite matrix, then

$$\sqrt{n}(\hat{\beta} - \beta) \to_d N_p(0, \sigma^2 V^{-1})$$

[This is a variant of the result we established in 581 via the Lindeberg - Feller CLT.]
(d) Find the mean and variance of Z_n.
(e) Suppose that:
 (i) $n^{-1}\mathbf{X}^T\mathbf{X} \to V$, a positive definite matrix;
 (ii) $\mathbf{X}^T\mathbf{1}/n \to \mathbf{h}$ with $\mathbf{h}^T V^{-1}\mathbf{h} < 1$;
 (iii) $\max_{1\le i\le n} h_{ii} \to 0$ where h_{ii} are the diagonal elements of the hat matrix H.

Show that if (i) - (iii) hold, then the bootstrap fails in the sense that the random variable Z_n in (b) converges in distribution to a proper random variable rather than to zero.
Hint: show that (iii) implies that $\max_{1\le i\le n}|c_{ni}| \to 0$ where $\mathbf{c} = n^{-1/2}(I - H)\mathbf{1}$.

6. **Optional bonus problem 2:** Suppose now that the bootstrap residuals are drawn from the collection of *centered* residuals $\hat{\epsilon} - \mathbf{1}(\mathbf{1}^T\hat{\epsilon}/\mathbf{n})$ Compute $E_*(\sqrt{n}(\hat{\beta}^* - \hat{\beta}))$ and $E_*(\sqrt{n}(\hat{\beta}^* - \hat{\beta}))^{\otimes 2}$ for this bootstrap resampling scheme.