

## Statistics 523, Problem Set 4

Wellner; 4/19/2017

**Reading:** Shorack, PfS Course Notes, Chapter 7, Examples 7.5.2 and 7.5.3;  
Shorack, PfS Course Notes, Appendix, pages 426 - 428;  
Shorack, PfS Course Notes, Chapter 11, pages 281-287;  
Chen, Goldstein, & Shao, chapter 2.

**Reminder:** Make-up lecture 2: Wednesday, April 26, 12:30 - 1:20, CMU 230.

**Due:** Wednesday, April 26, 2017.

1. Exercise 7.5.4, PfS, page 146: Suppose that  $X_1, \dots, X_n$  are i.i.d. with continuous distribution function  $F$ . Show that with probability 1 all the observations are distinct. [Hint: use corollary 2 to Theorem 5.1.3.]
2. Suppose that  $\{b_i\}_{i=1}^N$  and  $\{c_i\}_{i=1}^N$  are two sequences of real numbers, and write  $c(i) \equiv c_i$ . Suppose that  $\underline{R} = (R_1, \dots, R_N)$  is distributed uniformly over the collection  $\Pi_N$  the collection of all permutations of  $\{1, \dots, N\}$ ; i.e.  $P(\underline{R} = \underline{r}) = 1/N!$  for all  $\underline{r} \in \Pi_N$ . Let  $S \equiv S_N \equiv \sum_{j=1}^N b_j c(R_j)$ . Show that  $Var(S) = (N-1)^{-1} B_N^2 \cdot C_N^2$  where  $B_N^2 = \sum_{j=1}^N (b_j - \bar{b}_N)^2$  and  $C_N^2 = \sum_{j=1}^N (c_j - \bar{c}_N)^2$ .
3. Suppose that  $X_1, \dots, X_n$  are the numbers resulting from sampling without replacement from an urn consisting of balls with the numbers  $a_1, \dots, a_N$  on the  $N$  balls. Let  $\bar{a}_N \equiv \bar{a} \equiv N^{-1} \sum_{i=1}^N a_i$  and  $\sigma_a^2 \equiv N^{-1} \sum_{i=1}^N (a_i - \bar{a})^2$ . Let  $T_n \equiv X_1 + \dots + X_n$ . Verify that for  $j \neq k, j, k \in \{1, \dots, N\}$ ,

$$Cov[X_j, X_k] = Cov[X_1, X_2] = -\frac{\sigma_a^2}{N-1}$$

and that

$$Var(T_n/n) = \frac{\sigma_a^2}{n} \left(1 - \frac{n-1}{N-1}\right).$$

The factor  $(1 - (n-1)/(N-1))$  is sometimes called the *finite-sampling correction factor*; note that the variance of the mean is *smaller* than the variance of the mean under sampling with replacement (namely  $n^{-1}\sigma_a^2$ ).

4. Suppose that  $Y_1, Y_2, \dots$  are i.i.d. with distribution function  $G$  and characteristic function  $\varphi(t) = E \exp(itY_1)$ . Let  $N_\lambda$  a random variable with Poisson( $\lambda$ ) distribution and assume that  $N_\lambda$  is independent of the  $\{Y_i\}$ 's. Let  $S \equiv S_\lambda \equiv \sum_{j=1}^{N_\lambda} Y_j$ . Find the characteristic function  $\phi_S$  of  $S$ .

5. **Bonus problem 1: Beyond zero-bias Stein identities:** There are several other ways of using the Stein identity (or characterization of the normal distribution) to prove central limit theorems and establish rates of convergence. Some of the most important of these are explained in Chen, Goldstein, and Shao (2011), chapter 2:

(a) For sums of independent random variables: the  $K$  function approach:

(b) Exchangeable pairs.

(c) Size biased transformation.

For your choice of one of (a), (b), or (c), explain the basic identity analogous to the identity  $E[Wf(W)] = \sigma^2 E[f'(W^*)]$  used in the zero-biased distribution approach.