# Exercises 3, Stat 206B/Math 223B, Spring 2011

1. Let $(W, W')$ an $a$-Stein pair with variance $\sigma^2$; that is $(W, W')$ is exchangeable and
$$\mathbb{E}[W' - W | W] = -aW.$$
Let $dF(w, w')$ denote the distribution of $(W, W')$.

   (a) Show that the measure $G$ defined by
$$dG(w_1, w_2) = \frac{(w_1 - w_2)^2}{2a\sigma^2} dF(w_1, w_2)$$
is a probability measure. We refer to random variables with this measure as "square-biased."

   (b) Let $(W_1, W_2)$ have distribution $dG(w_1, w_2)$, and let $U$ be a uniform $(0, 1)$ random variable independent of all else. Show that the random variable
$$W^* := UW_1 + (1 - U)W_2$$
has the zero-biased distribution of $W$.

2. Let $(a_{ij})_{i,j=1}^n$ be an $n \times n$ matrix, $\pi$ a random permutation of $\{1, \ldots, n\}$, and define $W = \sum_{i=1}^n a_{i\pi_i}$. The purpose of this exercise is to use Exercise 1 above and the zero-bias normal approximation theorem from class obtain an error in the approximation of $(W - \mathbb{E}[W])/\sqrt{\mathrm{Var}(W)}$ by a standard normal distribution.

   (a) Define $\tilde{a}_{ij} = a_{ij} - a_{\cdot j} - a_{i\cdot} + a_{\cdot\cdot}$ where
$$a_{\cdot j} = \frac{1}{n} \sum_i a_{ij}, \quad a_{i\cdot} = \frac{1}{n} \sum_j a_{ij}, \quad a_{\cdot\cdot} = \frac{1}{n^2} \sum_{i,j} a_{ij},$$
and let $\tilde{W} = \sum_i \tilde{a}_{i\pi_i}$. Show
$$\sum_i \tilde{a}_{ij} = \sum_j \tilde{a}_{ij} = 0, \tag{1}$$
and that assuming
$$\frac{1}{n-1} \sum_{i,j} \tilde{a}_{ij}^2 > 0, \tag{2}$$

we have

$$\frac{\tilde{W} - \mathbb{E}[\tilde{W}]}{\sqrt{\mathrm{Var}(\tilde{W})}} = \frac{W - \mathbb{E}[W]}{\sqrt{\mathrm{Var}(W)}}.$$

Show that these facts imply that no generality is lost in proving a CLT for $W$ assuming (1) and (2) with $\tilde{a}_{ij}$ replaced by $a_{ij}$, and also

$$\frac{1}{n-1}\sum_{i,j} a_{ij}^2 = 1.$$

Make this assumption for the remainder of this exercise.

(b) Let $\tau$ be a random transposition of $\{1, \ldots, n\}$; that is $\tau_{I,J} := (IJ)$, where $(I, J)$ is chosen uniformly among all $n(n-1)$ ordered pairs of $\{1, \ldots, n\}$. If $\pi' = \pi\tau_{I,J}$ and $W' = \sum_i a_{i\pi'_i}$, then show that $(W, W')$ is a $2/(n-1)$ Stein pair.

(c) Let the random indices $I_1$, $J_1$, $K_1$, $L_1$ be chosen proportional to the squared difference of $(W - W')^2$; that is

$$\mathbb{P}(I_1 = i, J_1 = j, K_1 = k, L_1 = l) = c(a_{ik} + a_{jl} - a_{il} - a_{jk})^2,$$

independent of $\pi$. Define the pair $(\pi^{(1)}, \pi^{(2)})$ by

$$\pi^{(1)} = \begin{cases} \pi\tau_{\pi^{-1}(K_1),J_1}, & L_1 = \pi(I_1), K_1 \neq \pi(J_1) \\ \pi\tau_{\pi^{-1}(L_1),I_1}, & L_1 \neq \pi(I_1), K_1 = \pi(J_1) \\ \pi\tau_{\pi^{-1}(K_1),I_1}\tau_{\pi^{-1}(L_1),J_1}, & \text{otherwise,} \end{cases}$$

and $\pi^{(2)} = \pi^{(1)}\tau_{I_1,J_1}$. It is a fact that $(W_1, W_2) = (\sum_i a_{i\pi_i^{(1)}}, \sum_i a_{i\pi_i^{(2)}})$ has the square biased distribution of $(W, W')$; assume this to be true. Show that there are variables $S$, $T$, $T_1$, and $T_2$ satisfying

$$W = S + T, \quad W_1 = S + T_1, \quad W_2 = S + T_2,$$

and such that $|T|, |T_1|$, and $|T_2|$ are all bounded by $4\max_{i,j}|a_{ij}|$, so that

$$|W - W^*| \leqslant 8 \max_{i,j} |a_{ij}|.$$

Use this fact to obtain an error in the approximation of $W$ by a standard normal distribution.