STAT 701 HW2 Solutions, 2/27/23

(1). (#5.2.4) (a). Write the desired function as $\gamma(\rho)$, and use the fact that U_1 and $Z_1 = (V_1 - \rho U_1)/\sqrt{1-\rho^2}$ are independent $\mathcal{N}(0,1)$ random variables. Then

$$\gamma(\rho) \;=\; \int_0^\infty \int_{-\rho u_1/\sqrt{1-\rho^2}}^\infty \phi(u_1)\,\phi(z_1)\,dz_1\,du_1 \;=\; \int_0^\infty \phi(u_1)\,\Phi(-\rho u_1/\sqrt{1-\rho^2})\,du_1$$

Then $\gamma(0) = 1/4$ and

$$\gamma'(\rho) \;=\; \frac{-1}{2\pi} \int_0^\infty \frac{u_1}{(1-\rho^2)^{1/2}}\, e^{-u_1^2/(2(1-\rho^2))}\,du_1 \;=\; \frac{-1}{2\pi}\,(1-\rho^2)^{1/2}$$

and the desired formula uis the unique solution of this ordinary differential equation.

(b). This is an obvious conclusion from the Law of Large Numbers, given our class discussion of Generalized Method of Moments estimators, if $\bar{X}, \bar{Y}$ are respectively replaced by $E(X_1), E(Y_1)$ But it is easy to argue directly (whether (X_i, Y_i) are bivariate-normal or not, that

$$\frac{1}{n}\sum_{i=1}^n \left[I_{[X_i>\bar{X},\,Y_i>\bar{Y}]} \;-\; I_{[X_i>E(X_1),\,Y_i>E(Y_1)]}\right] \;\to\; 0 \quad \text{as} \quad n\to\infty$$

(c). However, when the joint distribution of X_i, Y_i is not normal, then the formula in (a) is no longer valid, so the estmator in (b) is consistent for $\sin(2\pi P(X_1 > E(X_1),\, Y_1 > E(Y_1)))$.

(2). (#5.3.10) Example 5.3.6 showed that for bivariate-normal data, the sample correlation coefficient ρ^* is asymptotically normal with mean ρ and variance ($1/n$ times) $(1-\rho^2)^2$. Since the trnasformation $g(x) = (1/2)\log((1+x)/(1-x))$ has $g'(x) = (1/2)(1/(1+x) + 1/(1-x)) = 1/(1-x^2)$, the Delta Method shows that

$$\sqrt{n}(g(\rho^*) - g(\rho)) \;\overset{\mathcal{D}}{\to}\; \mathcal{N}(0,\,1) \qquad \text{as} \qquad n\to\infty$$

(3). (#5.3.33) (a). Calculation of these MLEs is completely sttraightforward. (b). Since $\hat{\mu}_i$ is the average of the normal observations in the i'th row of the matrix X_{ij} and $s_i^2 = (k-1)^-1]sum_{j=1}^k (X_{ij} - \hat{\mu}_i)^2$ is the sample variance of the i'th row, we know that $(k-1)\hat{s}_i^2/\sigma^2 \sim \chi^2_{k-1}$. Since the s_i^2 are *iid* across i, the Law of Large Numbers implies

$$\hat{\sigma}^2 = \frac{k-1}{k}\cdot\frac{1}{p}\sum_{i=1}^p s_i^2 \to \frac{k-1}{k}E(s_1^2 \;=\; \frac{k-1}{k}\cdot\sigma^2$$

(c) So $\hat{\sigma}^2$ is inconsistent for σ^2, but $k\hat{\sigma}^2/(k-1)$ is consistent.

(4). (#5.4.1) (a). Consider the function $g(t) = E_P(\psi(X_1 - t)))$. It follows immediately from the assumed properties of ψ that $|g(t)| \le M$ and $g(t)$ is non-increasing as a function of t. In addition, from the bounded convergence theorem we find that for a sequence $t_n \nearrow \infty$,

$$g(t_n) \;=\; E_P(\psi(X_1 - t_n)) \;\to\; E_P(\psi(-\infty)) = \psi(-\infty) < 0$$

and similarly if $t_n \nearrow t < \infty$, also $g(t-) = g(t)$, and based on decreasing sequences, $g(-\infty) = \psi(\infty) > 0$ and $g(t-)$ exists. That is, the function g is bounded nondecreasing with limits from the left and right, and $g(-\infty) > 0 > g(\infty)$. (Note that the existence of limits from the left [respectively, from the right] does not imply that $g(t-) = g(t)$ [resp. that $g(t+) = g(t)$].) Therefore $\theta(P)$ exists and can be any value in the interval between $\sup\{t : \; g(t) > 0\}$ and $\inf\{t : \; g(t) < 0$.

(b). Now we are assuming $\theta(P)$ is unique, i.e. that $g(\theta(P)-) \ge 0 \ge g(\theta(P)+)$, and for arbitrary $\epsilon > 0$, $g(\theta(P) - \epsilon) > 0 > g(\theta(P) + \epsilon)$. The Weak Law of Large Numbers says for $\delta = \min(\theta(P) - \epsilon) > 0 > g(\theta(P) + \epsilon)) > 0$,

$$P\Big(n^{-1} \sum_{i=1}^{n} \psi(X_i - \theta(P) + \epsilon) < -\delta/2 \quad \text{or} \quad n^{-1} \sum_{i=1}^{n} \psi(X_i - \theta(P) - \epsilon) > \delta/2\Big) \;\to\; 0$$

It follows from this that, even though $\hat{\theta}_n = \theta(\hat{P})$ need not be unique, $P(|\hat{\theta}_n| \ge \epsilon) \to 0$ as $n \to \infty$. Since $\epsilon > 0$ is arbitrariy small, this proves consistency.

(c). This part involves a Taylor's series exercise using differentiability of g only at the point $\theta = \theta(P)$:

$$g(\theta + t/\sqrt{n}) \equiv E(\psi(X_1 - \theta - t/\sqrt{n})) = g(\theta) - t\, g'(\theta)/\sqrt{n} + o(1/\sqrt{n}) \qquad (*)$$

From this, we conclude that the right- and left-hand limits at θ satisfy $g(\theta+) = g(\theta-) = 0$, and therefore also $g(\theta)$ (which is sandwiched between these) is 0.

Now we use the hint:

$$P(\sqrt{n}\,(\hat{\theta}_n - \theta) < t) \;=\; P(n^{-1/2} \sum_{i=1}^{n} (\psi(X_i - \theta_n) - g(\theta_n)) \;>\; -\sqrt{n} g(\theta_n))$$

Since we saw above in (*) that $\sqrt{n}\, g(\theta_n) \;\to\; -t\, g'(\theta)$, the first normal-convergence assumption in (c) implies

$$P(\sqrt{n}\,(\hat{\theta}_n - \theta) < t) \;\to\; 1 - \Phi(t g'(\theta)/\tau(\theta)) \;=\; \Phi(-t g'(\theta)/\tau(\theta))$$

which is exactly the desired normal-convergence assertion.

(d). This part has a misprint. (A6) expressed the equality between $E(-\psi'(X_1 - \theta))$ and $-\mathrm{Cov}(\psi(X_1-\theta), \partial \log f(X_1,\theta)/\partial\theta)$. However, even with this change in the covariance formula, we need an additional Dominated-Convergence assumption to make the derivative of g at θ equal to $E(-\psi'(X_i - \theta))$.

(A). Let $X \sim \chi^2_m = \text{Gamma}(m/2, 1/2)$ and $Y \sim \chi^2_n = \text{Gamma}(n/2, 1/2)$ be independent. Then $W = X/(X+Y) \sim \text{Beta}(m/2.n/2)$ takes values in $(0,1)$, and $R = (n/m)X/Y = (n/m)W/(1-W) \sim F_{m,n}$. Therefore with $b(w) = nw/(m(1-w))$ and $(b^{-1})'(r) = mn/(n+mr)^2$, the $F_{m,n}$ density is given by

$$f_R(r) \;=\; \frac{\Gamma((m+n)/2)}{\Gamma(m/2)\,\Gamma(n/2)}\,(\frac{mr}{n+mr})^{m/2-1}\,(\frac{n}{n+mr})^{n/2-1}\,\frac{mn}{(n+mr)^2}$$

For the t_k density we note that, this random variable T is the square root of an $F_{1,k-1}$ density, symmetrized by a random $\pm$ sign. So by one more transformation, with $R \sim F_{1,k-1}$, we find for all real t,

$$f_T(t) \;=\; \frac{1}{2} f_{\sqrt{R}}(t^2) \;=\; \frac{1}{2}\cdot\frac{\Gamma(k/2)}{\sqrt{\pi}\cdot\Gamma((k-1)/2)}(k-1)^{(k-1)/2}\,(t^2)^{-1/2}\,(k-1+t^2)^{-k/2}\,(2t)$$

(B). In (a), let U be any orthogonal transformation with first column $\mathbf{1}/\sqrt{n}$, and Z be a column vector of n *iid* $\mathcal{N}(0,1)$ random variables, and $W = U(Z + \mu/\sqrt{n})$. Then $W_1 = \sqrt{n}\cdot\bar{Z} + \mu$, and W_j for $j \geq 2$ are $\mathcal{N}(0,1)$ random variables independent of one another and of $W_1 \sim \mathcal{N}(\mu, 1)$, and $\sum_{j=1}^n (Z_j + \mu/\sqrt{n})^2 = W_1^2 + \sum_{j=2}^n W_j^2$, as desired.

In (b), the idea is similar, but the first column of the orthogonal matrix U must now be $\mathbf{v}/\|\mathbf{v}\|$. (The other columns are any other elements of an orthonormal basis of the subspace of $\mathbb{R}^n$ orthogonal to $\mathbf{v}$. Now W has all entries independent, with $W_1 = \mathbf{v}^{tr} Z + \|v\| \sim \mathcal{N}(0, \|\mathbf{v}\|^2)$, and the remaining $W_j \sim \mathcal{N}(0,1)$. So $\sum_{j=1}^n (Z_j + v_j)^2 = W_1^2 + \sum_{j=2}^n W_j^2$ is noncentral chi-squared with noncentrality parameter $\|\mathbf{v}\|^2$.