STAT 701 HW1 Solutions, 2/10/23

(1). (#4.4.14) We wanted exact CIs for this problem.

(a) For μ (with σ^2 unknown) : $\bar{X} \pm t_{n-1,\alpha/2}\, S/\sqrt{n}$.

(b) For σ (with μ unknown) : $[\{(n-1)\, S^2/F^{-1}_{\chi^2_{n-1}}(1-\alpha/2)\}^{1/2},\ \{(n-1)\, S^2/F^{-1}_{\chi^2_{n-1}}(\alpha/2)\}^{1/2}]$.

(c) A (non-rectangular) confidence region with level $1-\alpha$ for (μ, σ) is defined as follows: let $\alpha^* = 1 - \sqrt{1-\alpha}$ and

$$c_1 = z_{\alpha^*/2}, \qquad c_2 = \{(n-1)\, S^2/F^{-1}_{\chi^2_{n-1}}(1-\alpha^*/2)\}^{1/2}, \qquad c_3 = \{(n-1)\, S^2/F^{-1}_{\chi^2_{n-1}}(1-\alpha^*/2)\}^{1/2}$$

Then the desired region is

$$\{(\mu, \sigma) : \quad \bar{X} - c_1\, \sigma/\sqrt{n} \le \mu \le \bar{X} + c_1\, \sigma/\sqrt{n}, \qquad c_2 \le \sigma \le c_3 \quad \}$$

(2). (#4.5.3) Problem 4.5.2 referred to in this problem is just a one-sided version of the idea we expressed in class that a Confidence Region $CI(\alpha \underline{X})$ of level $1-\alpha$ automatically provides a size α hypothesis test for $H_0 : \theta = \theta_0$ regarding the unknown parameter θ in a statistical problem based on data $\underline{X}$. The rejection region for that test is $\{\underline{X} : \ \theta_0 \notin CI(\alpha, \underline{X}\}$. That problem also asked you to note that in the one-sided setting $CI(\alpha, \underline{X}) = (-\infty, UCB(\alpha, \underline{X})]$ the same test has level α (i.e., size $\le \alpha$) for $H_0^* : \ \theta \ge \theta_0$

In this problem, we consider the upper confidence bound for unknown σ^2 based on normal data given by (4.4.2) in the book, with $\alpha_1 = \alpha, \alpha_2 = 0$, that is, the UCB is $(n-1)\, S^2/F^{-1}_{\chi^2_{n-1}}(\alpha)$, so the desired test in (a) rejects $H_0^* \ : \ \sigma^2 \ge \sigma_0^2$ when $(n-1)\, S^2/F^{-1}_{\chi^2_{n-1}}(\alpha) < \sigma_0^2$.

It remains to find the power $K(\sigma^2)$ of the test in (a), at values $\sigma^2 < \sigma_0^2$. That is, if $W = (n-1)S^2/\sigma^2$ denotes a χ^2_{n-1} distributed random variable under $P_{\sigma^2}(\cdot)$, then

$$K(\sigma^2) \ = \ P_{\sigma^2}((n-1)\, S^2/F^{-1}_{\chi^2_{n-1}}(\alpha) < \sigma_0^2) \ = \ F_{\chi^2_{n-1}}(F^{-1}_{\chi^2_{n-1}}(\alpha) \cdot (\sigma_0^2/\sigma^2))$$

(3). (#5.3.8(a)-(c)) (a) The likelihood in the normal-data case written in terms of sufficient statistics is

$$C - (n_1/2) \log(\sigma_1^2) - n_2/2 \log(\sigma_2^2) - \frac{1}{2\sigma_1^2}\{(n_1-1)S_1^2 + n_1(\bar{X}-\mu_1)^2\} - \frac{1}{2\sigma_2^2}\{(n_2-1)S_2^2 + n_2(\bar{X}-\mu_2)^2\}$$

Regardless of assumptions about the variance, the MLE's for μ_1, μ_2 are respectively $\bar{X}, \bar{Y}$. So the maximized log-Likelihood under the general hypothesis, with $\hat{\sigma}_j^2 = (1 - 1/n_j)\, S_j^2$, is

$$C - \frac{n_1}{2} \log(1 - 1/n_1) - \frac{n_2}{2} \log(1 - 1/n_2) - \frac{n_1 + n_2}{2} - \frac{n_1}{2} \log S_1^2 - \frac{n_2}{2} \log S_2^2$$

The maximum likelihood of σ^2 under the hypothesis $H : \sigma_1^2 = \sigma_2^2 = \sigma^2$ is the pooled MLE $\hat{\sigma}^2 = (n_1 + n_2)^{-1}((n_1 - 1)S_1^2 + (n_2 - 1)S_2^2)$. From that it is easy to check that the log of the Likelihood Ratio Statistic is

$$-\frac{n_1 + n_2}{2}\log(\hat{\sigma}^2 + \frac{n_1}{2}\log S_1^2 + \frac{n_2}{2}\log S_2^2$$

which is a monotone function of S_1^2/S_2^2.

(b). Since $(n_k - 1)S_k^2/\sigma_k^2$ are independent $\chi^2_{n_k-1}$ random variables for $k = 1, 2$, by definition of the F_{n_1,n_2} distribution the ratio $(S_1^2/\sigma_1^2)/(S_2^2/\sigma_2^2)$ has this distribution.

(c). Now we assume X_i has mean μ_1 and variance σ_1^2 and $Var(Y_j) = \sigma_1^2$ while Y_j is equal in distribution to $(X_i - a)/b$. Clearly $b = 1$, $a = \mu_1 - \mu_2$, and $\kappa = \mathrm{Var}((X_1 - \mu_1)/\sigma_1 = \mathrm{Var}((Y_1 - \mu_2)/\sigma_1))$. We also assume that the ratio $n_2/n_1 = \lambda$ is fixed as the sample sizes go to ∞. Then it is easy to check that,

$$\sqrt{n_1}\,(S_1^2 - \sigma_1^2) \overset{P}{\approx} \sqrt{n_1}\,\Big[\frac{1}{n_1}\sum_{i=1}^{n_1}(X_1 - \mu_1)^2 - \sigma_1^2\Big] \overset{\mathcal{D}}{\to} \mathcal{N}(\sigma_1^2,\, \kappa\sigma_1^4)$$

and a similar limit theorem holds (with the same distributional limit) for $\sqrt{n_2}\,(S_2^2 - \sigma_1^2)$. Therefore the Delta Method applied, with function $g(x, y) = (\sigma_1^2 + x)/(\sigma_1^2 + y) - 1$, to

$$\sqrt{n_1}(S_1^2/S_2^2 - 1) \;=\; \sqrt{n_1}\Big[\frac{\sigma_1^2 + (S_1^2 - \sigma_1^2)}{\sigma_1^2 + (S_2^2 - \sigma_2^2} \;-\; 1\Big]$$

shows that this displayed expression is asymptotically equal to

$$\sqrt{n_1}\,\frac{S_1^2 - \sigma_1^2}{\sigma_1^2} \;-\; \frac{1}{\sqrt{\lambda}}\sqrt{n_2}\,\frac{S_2^2 - \sigma_1^2}{\sigma_1^2} \overset{\mathcal{D}}{\to} \mathcal{N}(0,\, \kappa\,(1 + 1/\lambda))$$

It follows that $P(\sqrt{n_1}\,(S_1^2/S_2^2 - 1) \le \sqrt{\kappa(n_1 + n_2)/n_2}\,\Phi^{-1}(1 - \alpha)) \;\to\; 1 - \alpha$, which is what we were asked to prove.

(4). (#5.3.15(b)) This is a slightly disguised version of a Delta Method problem. Let $S_n = \sum_{i=1}^n Z_i^2 \sim \chi^2_n$, which has expectation n, and apply the Delta Method Taylor expansion idea to $S_n/n)^{1/3}$. Then

$$\sqrt{n}(\,\left(\frac{S_n}{n}\right)^{1/3} - 1) \;\approx\; \frac{1}{3}\,(S_n/n - 1) \;+\; O_P(n^{-1/2})$$

and (5.3.6) tells us that $\mathrm{Var}(S_n/n) \;=\; (1/3)^2\,(2/n) \;+\; O(n^{-3/2})$. Therefore,

$$P(S_n \le x) \;=\; P(\sqrt{n}\,((S_n/n)^{1/3} - 1) \;\le\; \sqrt{n}\,(x^{1/3} - 1)) \;\approx\; \Phi(\sqrt{n}\,(x^{1/3} - 1)\,/\,\sqrt{2/9}$$

and we make the approximation slightly better by adding an extra term $2/(9n)$ to the centering for $(S_n/n)^{1/3}$ using a second Taylor series term, since

$$E((S_n/n)^{1/3}) \approx E\Big[1 + (1/3)(S_n/n - 1) + (1/2)(1/3)(-2/3)(S_n/n - 1)^2\Big] \;=\; 1 - 2/(9n)$$

(A). The main idea here is to use the identity that the probability mass function with parameter a identically sums to 1. The Method of Moments in this problem is based on the moment formulas

$$E(W_1) = \sum_{k=1}^{\infty} \frac{k^{\alpha+1}\, e^{-k}}{C(\alpha,0)} \;=\; \frac{C(\alpha+1,0)}{C(\alpha,0)} \quad , \qquad E(W_1^2) = \sum_{k=1}^{\infty} \frac{k^{\alpha+2}\, e^{-k}}{C(\alpha,0)} \;=\; \frac{C(\alpha+2,0)}{C(a,0)}$$

and $\mathrm{Var}(W_1) \;=\; (C(\alpha,0))^{-2}\,(C(\alpha+2,0)\,C(\alpha,0) - C(\alpha+1,0)^2)$.

The canonical exponential family representation $p(w,\alpha) \;=\; e^{-k}\cdot \exp(\alpha\log(k) - \log(C(\alpha,0))$, i.e. in the exponential family, $A(\alpha) \equiv \log(C(\alpha,0)$. The derivatives of $A(\cdot)$ are given by

$$A'(\alpha) \;=\; C(\alpha,1)/C(\alpha,0) \quad , \qquad A''(\alpha) \;=\; C(\alpha,2)/C(\alpha,0) - (C(\alpha,1)/C(\alpha,0))^2$$

and the per-observation Fisher Information about α is $I(\alpha) = A''(\alpha)$.

If the function g is the inverse function of $C(a{+}1,0)/C(a,0)$, i.e., $g(r)$ is the solution of $C(a{+}1,0)/C(a,0) = r$, then the method of moment estimator of α is $\tilde{\alpha} = g(\bar{W})$, and $\sqrt{n}\,(g(\bar{W}) - \alpha)$ is asymptotically normally distributed with mean 0 and variance given according to the Delta Method by $(g'(\rho))^2\cdot \mathrm{Var}(W_1)$, where $\rho = C(\alpha+1,0)/C(\alpha,0)$. Since

$$g'(\rho) = 1/\frac{d}{d\alpha}\{C(\alpha{+}1,0)/C(\alpha,0)\} = (C(\alpha{+}1,1)C(\alpha,0) {-} C(\alpha{+}1,0)C(\alpha,1))^{-1}(C(\alpha,0))^2$$

we conclude

$$\mathrm{a.var}(\tilde{\alpha}) \;=\; (C(\alpha,0))^2\, \frac{C(\alpha+2,0)\,C(\alpha,0) - C(\alpha+1,0)^2}{(C(\alpha+1,1)C(\alpha,0) - C(\alpha+1,0)C(\alpha,1))^2}$$

and the Asymptotic Relative Efficiency is $1/A''(\alpha)$ divided by $\mathrm{a.var}(\tilde{\alpha})$.

(B). Let $T_1 = n^{-1}\sum_{i=1}^n (X_i{+}Y_i)^2$, $T_2 = n^{-1}\sum_{i=1}^n (X_i{-}Y_i)^2$. Then $E(T_1{+}T_2)/4 = \sigma_1^2$, $E(T_1 - T_2)/4 = \rho\sigma^2$. So the generalized method of moments estimator of $\rho\sigma^2$ is

$(T_1 - T_2)/4 = n^{-1}\sum_{i=1}^n X_i\,Y_i$, and its asymptotic variance (after stripping the factor $1/n$) is $\text{Var}(X_1Y_1) =$

$$E(X_1^2Y_1^2) - \rho^2\sigma^4 = E(X_1^2(Y_1 - \rho X_1 + \rho X_1)^2) - \rho^2\sigma^4 \;=\; \sigma^4(1-2\rho^2) + \rho^2\,E(X_1^4) \;=\; (1+\rho^2)\sigma^4$$

This is to be compared to the inverse Fisher information, based on the log-density (for a single observation (X_i, Y_i)

$$-\log(2\pi) - \frac{1}{2}\log((1-\rho^2)\sigma^4) \;-\; (x^2 + y^2 - 2\rho xy)/(2\sigma^2(1-\rho^2))$$

The (per-observation) Fisher Information matrix for (σ^2, ρ) is

$$\begin{pmatrix} \sigma^{-4} & \rho/((1-\rho^2)\sigma^2) \\ \rho/((1-\rho^2)\sigma^2) & (3+5\rho^2)/(1-\rho^2)^2 \end{pmatrix}$$

The inverse Fisher Information is

$$\frac{1}{3+4\rho^2}\begin{pmatrix} \sigma^4(3+5\rho^2) & -\rho\sigma^2(1-\rho^2) \\ -\rho\sigma^2(1-\rho^2) & (1-\rho^2)^2 \end{pmatrix}$$

and the Cramer-Rao lower bound for the function $\psi(\sigma^2, \rho) \;=\; \rho\sigma^2$ is obtained by pre- and post-multiplying the last matrix by $\nabla\psi \;=\; (\rho, \sigma^2)$. So the Cramer-Rao bound is

$$\frac{\sigma^4}{3+3\rho^2}\left[\rho^2(3+5\rho^2) - 2\rho^2(1-\rho^2) + (1-\rho^2)^2\right] \;=\; \frac{\sigma^4\,(1-\rho^2+8\rho^4)}{3+4\,\rho^2}$$

and the relative efficiency is

$$\frac{1-\rho^2+8\rho^4}{(3+4\rho^2)(1+\rho^2)}$$