Jiapeng Wang,

5.8 Let $X_1, \dots, X_n$ be a random sample, where $\bar{X}$ and S^2 are calculated in the usual way.

(a) Show that

$$S^2 = \frac{1}{2n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} (X_i - X_j)^2.$$

Assume now that the X_is have a finite fourth moment, and denote $\theta_1 = \mathrm{E}X_i, \theta_j = \mathrm{E}(X_i - \theta_1)^j, j = 2, 3, 4$.

(b) Show that $\mathrm{Var}\, S^2 = \frac{1}{n}(\theta_4 - \frac{n-3}{n-1}\theta_2^2)$.

(c) Find $\mathrm{Cov}(\bar{X}, S^2)$ in terms of $\theta_1, \dots, \theta_4$. Under what conditions is $\mathrm{Cov}(\bar{X}, S^2) = 0$?

(a). $\frac{1}{2n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} (X_i - X_j)^2 = \frac{1}{2n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} (X_i - \bar{X} + \bar{X} - X_j)^2$

$= \frac{1}{2n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} [(X_i - \bar{X})^2 - 2(X_i - \bar{X})\cdot(X_j - \bar{X}) + (X_j - \bar{X})^2]$.

$= \frac{1}{2n(n-1)} [\sum_{i=1}^{n} n(X_i - \bar{X})^2 - 2\sum_{i=1}^{n}(X_i - \bar{X}) \sum_{j=1}^{n}(X_j - \bar{X}) + n\sum_{j=1}^{n}(X_j - \bar{X})^2]$.

$= \frac{n}{2n(n-1)} \sum_{i=1}^{n} (X_i - \bar{X})^2 + \frac{n}{2n(n-1)} \sum_{j=1}^{n}(X_j - \bar{X})^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = S^2$

(b). Without loss of generality, suppose $E(X_i) = 0$, $Var(X_i) = \sigma^2$ $\theta_1 = 0$

so $\theta_2 = E(X_i - E(X_i))^2 = E(X_i^2) = Var(X_i) = \sigma^2$.

Now we consider $(n-1)^2 E((S^2)^2) = E((\sum_{i=1}^{n} X_i^2 - n\bar{X}^2)^2)$

$= E(\sum_{i=1}^{n} X_i^2)^2 - 2nE(\sum_{i=1}^{n} X_i^2 \bar{X}^2) + n^2 E(\bar{X}^4)$

symmetry $\quad E(\sum_{i=1}^{n} X_i^2) - 2n^2 E(X_1^2 \bar{X}^2) + n^2 E(X_1 \bar{X}^3)$

$= E(\sum_{i=1}^{n} X_i^2)^2 - 2E(X_1^2(\sum_{i=1}^{n} X_i)^2) + \frac{1}{n}E(X_1(\sum X_i)^3)$

Now we calculate that $E(\sum X_i^2)^2 = E(\sum X_i^4 + 2\sum_{i<j} X_i^2 X_j^2)$

$= n\theta_4 + n(n-1)\sigma^4$

$E(X_1^2(\sum X_i)^2) = E(X_1^4 + \sum_{i \neq 1} X_1^2 X_i^2) = \theta_4 + (n-1)\sigma^4$

$$E(X_1(\sum X_i)^3) = E(X_1^4) + \binom{3}{1} E(X_1^2 \sum_{i \neq 1} X_i^2) = \theta_4 + 3(n-1)\sigma^4$$

Hence, $(n-1)^2 E((S^2)^2) = \frac{(n-1)^2}{n}\theta_4 + \left((n-1)^2 - \frac{n^2-4n+3}{n}\right)\sigma^4$

Finally, $Var(S^2) = E((S^2)^2) - E^2(S^2) = \frac{1}{n}\left(\theta_4 - \frac{n-3}{n-1}\sigma^4\right)$

$$= \frac{1}{n}\left(\theta_4 - \frac{n-3}{n-1}\theta_2^2\right)$$

(c). we could still assume $\theta_1 = 0$ Then

$$Cov(\bar{X}, S^2) = \frac{1}{2n^2(n-1)} E\left(\sum_{k=1}^{n} X_k \sum_{i=1}^{n}\sum_{j=1}^{n}(X_i - X_j)^2\right)$$

$$= \frac{2n(n-1)}{2n^2(n-1)} E X_i (X_i - X_j)^2 = \frac{1}{n}\theta_3$$ So when $Cov(\bar{X}, S^2) = 0$

$\theta_3 = 0$.

$\square$

5.11 Suppose $\bar{X}$ and S^2 are calculated from a random sample $X_1, \ldots, X_n$ drawn from a population with finite variance σ^2. We know that $\mathrm{E}S^2 = \sigma^2$. Prove that $\mathrm{E}S \leq \sigma$, and if $\sigma^2 > 0$, then $\mathrm{E}S < \sigma$.

5.11 We now consider $g(S) = S^2$, $g(S)$ is obviously a convex function,

From Jensen's inequality, we get $E\,g(S) = ES^2 \geq g(ES) = (ES)^2$

i.e $ES \leq \sigma$. if $\sigma^2 > 0$, then S is not a constant.

Thus, the inequality is strict, i.e $ES < \sigma$.

$\square$

5.15 Establish the following recursion relations for means and variances. Let $\bar{X}_n$ and S_n^2 be the mean and variance, respectively, of $X_1, \ldots, X_n$. Then suppose another observation, X_{n+1}, becomes available. Show that

(a) $\bar{X}_{n+1} = \frac{X_{n+1} + n\bar{X}_n}{n+1}$.

(b) $nS_{n+1}^2 = (n-1)S_n^2 + \left(\frac{n}{n+1}\right)(X_{n+1} - \bar{X}_n)^2$.

5.15. (a). $\bar{X}_{n+1} = \frac{\sum_{i=1}^{n+1} X_i}{n+1} = \frac{X_{n+1} + \sum_{i=1}^{n} X_i}{n+1} = \frac{X_{n+1} + n\bar{X}_n}{n+1}$

(b). $nS_{n+1}^2 = \sum_{i=1}^{n+1}(X_i - \bar{X}_{n+1})^2 = \sum_{i=1}^{n+1}\left(X_i - \frac{X_{n+1} + n\bar{X}_n}{n+1}\right)^2$

$= \sum_{i=1}^{n+1}\left(X_i - \frac{X_{n+1}}{n+1} - \frac{n\bar{X}_n}{n+1}\right)^2 = \sum_{i=1}^{n+1}\left[(X_i - \bar{X}_n) - \left(\frac{X_{n+1}}{n+1} - \frac{\bar{X}_n}{n+1}\right)\right]^2$

$= \sum_{i=1}^{n+1}\left[(X_i - \bar{X}_n)^2 + \frac{1}{(n+1)^2}(X_{n+1} - \bar{X}_n)^2 - \frac{2}{n+1}(X_i - \bar{X}_n)\cdot(X_{n+1} - \bar{X}_n)\right]$

$= \sum_{i=1}^{n}(X_i - \bar{X}_n)^2 + (X_{n+1} - \bar{X}_n)^2 + \frac{1}{n+1}(X_{n+1} - \bar{X}_n)^2 - \frac{2}{n+1}\sum_{i=1}^{n+1}(X_i - \bar{X}_n)\cdot(X_{n+1} - \bar{X}_n)$

$= (n-1)S_n^2 + \frac{n+2}{n+1}(X_{n+1} - \bar{X}_n)^2 - \frac{2}{n+1}(X_{n+1} - \bar{X}_n)\cdot(X_{n+1} - \bar{X}_n + n\bar{X}_n - n\bar{X}_n)$

$= (n-1)S_n^2 + \frac{n}{n+1}(X_{n+1} - \bar{X}_n)^2$

$\square$

$q/2$.

5.18 Let X be a random variable with a Student's t distribution with p degrees of freedom.

(a) Derive the mean and variance of X.

(b) Show that X^2 has an F distribution with 1 and p degrees of freedom.

(c) Let $f(x|p)$ denote the pdf of X. Show that

$$\lim_{p\to\infty} f(x|p) \to \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

at each value of x, $-\infty < x < \infty$. This correctly suggests that as $p \to \infty$, X converges in distribution to a $\mathrm{n}(0,1)$ random variable. (*Hint*: Use Stirling's Formula.)

(d) Use the results of parts (a) and (b) to argue that, as $p \to \infty$, X^2 converges in distribution to a χ_1^2 random variable.

(e) What might you conjecture about the distributional limit, as $p \to \infty$, of $qF_{q,p}$?

(a) $X \sim t_p$, we can write $X = \frac{U}{\sqrt{V/p}}$, where $U \sim n(0,1)$ and $V \sim \chi_p^2$, U and V are independent.

Then, $EX = E\frac{U}{\sqrt{V/p}} = EU \cdot E\left(\frac{1}{\sqrt{V/p}}\right)$, $EU=0$, $E\left(\frac{1}{\sqrt{V/p}}\right) < \infty$,

So, $EX=0$. $\quad \mathrm{Var}X = EX^2 - (EX)^2 = EX^2$.

Since $X^2 \sim F_{1,p}$ from Thm 5.3.8. Thus $\mathrm{Var}X = \frac{p}{p-2}$, $p>2$.

(b). $X^2 = \frac{U^2}{V/p}$, $\quad U \sim n(0,1) \Rightarrow U^2 \sim \chi_1^2$, Also since $V \sim \chi_p^2$ $\Rightarrow X^2 = \frac{U^2/1}{V/p} \sim F_{1,p}$

(c). We already learned that the pdf of X is

$$f(x|p) = \frac{\Gamma\left(\frac{p+1}{2}\right)}{\Gamma\left(\frac{p}{2}\right)} \frac{1}{(p\pi)^{\frac{1}{2}}} \frac{1}{\left(1+\frac{x^2}{p}\right)^{\frac{p+1}{2}}}$$

We first consider $\lim\limits_{p\to\infty} \left(1+\frac{x^2}{p}\right)^{-\frac{p+1}{2}} = \lim\limits_{p\to\infty} \left[\left(1+\frac{1}{\frac{p}{x^2}}\right)^{\frac{p}{x^2}}\right]^{-\frac{p+1}{2p}x^2}$

$$= e^{-\frac{x^2}{2}}, \quad x \neq 0.$$

Then: $\lim\limits_{p\to\infty} \frac{\Gamma\left(\frac{p+1}{2}\right)}{\Gamma\left(\frac{p}{2}\right)} \cdot \frac{1}{\sqrt{p\pi}} = \lim\limits_{p\to\infty} \frac{\sqrt{2\pi}\left(\frac{p-1}{2}\right)^{\frac{p-1}{2}+\frac{1}{2}} \cdot e^{-\frac{p-1}{2}}}{\sqrt{2\pi}\left(\frac{p-2}{2}\right)^{\frac{p-2}{2}+\frac{1}{2}} \cdot e^{-\frac{p-2}{2}}} \cdot \frac{1}{\sqrt{p\pi}}$

$$= \frac{e^{-\frac{1}{2}}}{\sqrt{\pi}} \cdot \lim_{p\to\infty} \frac{\left(\frac{p-1}{2}\right)^{\frac{p-1}{2}+\frac{1}{2}}}{\left(\frac{p-2}{2}\right)^{\frac{p-2}{2}+\frac{1}{2}} \cdot \sqrt{p}} = \frac{e^{-\frac{1}{2}} \cdot e^{\frac{1}{2}}}{\sqrt{\pi} \cdot \sqrt{2}} = \frac{1}{\sqrt{2\pi}}. \text{ (Stirling's Formula)}$$

(d). From (b), $X^2 \sim F_{1,p}$. Let $Y = X^2$.

Then, the pdf of Y can be denoted as.

$$f(y|p) = \frac{\Gamma\left(\frac{p+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{p}{2}\right)} \cdot \left(\frac{1}{p}\right)^{\frac{1}{2}} \frac{y^{-\frac{1}{2}}}{\left[1+\frac{y}{p}\right]^{\frac{(p+1)}{2}}}$$

Let $p \to \infty$, we can also use stirling's formula here.

$$\lim_{p\to\infty} f(y|p) = \left(\frac{p}{2}\right)^{\frac{1}{2}} \cdot \left(\frac{1}{p}\right)^{\frac{1}{2}} \cdot \frac{1}{\Gamma\left(\frac{1}{2}\right)} \cdot y^{-\frac{1}{2}} \cdot e^{-\frac{y}{2}}$$

$$= \frac{2^{-\frac{1}{2}}}{\Gamma\left(\frac{1}{2}\right)} \cdot y^{-\frac{1}{2}} \cdot e^{-\frac{y}{2}},$$ which is the pdf of χ^2_1 random variable

So, we get that X^2 converges in distribution to a χ^2_1 random variable.

(e). The random variable $qF_{q,p}$ could be thought as the sum of q random variables and each of them is the square of a random variable with a Student's t distribution. So the distribution limit of $qF_{q,p}$ is the sum of q random variables with χ^2_1 distribution, which is χ^2_q

$\square$

Problem A.0.1. Which of the following parametrizations are identifiable? Justify your answers.

1. Let $X_1, \ldots, X_p$ be independent with $X_i \sim \mathcal{N}(\alpha_i + \nu, \sigma^2)$. Write
$$\boldsymbol{\theta} = (\alpha_1, \alpha_2, \ldots, \alpha_p, \nu, \sigma^2),$$
and let $P_{\boldsymbol{\theta}}$ denote the distribution of $\boldsymbol{X} = (X_1, \ldots, X_p)$.
2. Same as (1) but with $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)$ restricted to the set
$$\left\{(a_1, \ldots, a_p) : \sum_{i=1}^{p} a_i = 0\right\}.$$
3. Let X and Y be independent $\mathcal{N}(\mu_1, \sigma^2)$ and $\mathcal{N}(\mu_2, \sigma^2)$, $\boldsymbol{\theta} = (\mu_1, \mu_2)$, and we observe $Y - X$.
4. Let X_{ij} for $i = 1, \ldots, p; j = 1, \ldots, b$ be independent with $X_{ij} \sim \mathcal{N}(\mu_{ij}, \sigma^2)$, where $\mu_{ij} = \nu + \alpha_i + \lambda_j$, $\boldsymbol{\theta} = (\alpha_1, \ldots, \alpha_p, \lambda_1, \ldots, \lambda_b, \nu, \sigma^2)$, and $P_{\boldsymbol{\theta}}$ is the distribution of $X_{11}, \ldots, X_{pb}$.
5. Same as (4) but with $\boldsymbol{\alpha}$ and $\boldsymbol{\lambda}$ satisfying the conditions $\sum_{i=1}^{p} \alpha_i = 0$ and $\sum_{j=1}^{b} \lambda_j = 0$.

A.0.1, 1. $X_i \sim N(\alpha_i + \nu, \sigma^2)$. So we get the pdf of $\vec{X}$: $P_\theta = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^p e^{\frac{-1}{2\sigma^2}\sum_{i=1}^{p}(X_i - \nu - \alpha_i)^2}$

We can find a counterexample, let $\alpha_1 = \alpha_2 = \cdots = \alpha_p = \nu$

Then. $P_\theta = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^p e^{\frac{-1}{2\sigma^2}\sum_{i=1}^{p} X_i^2}$. However, as α_i and ν can take different values, $P_{\theta_1} = P_{\theta_2}$ cannot imply $\theta_1 = \theta_2$.

2. We can set $\vec{\theta}_1 = (\alpha_1, \alpha_2, \cdots, \alpha_p, \nu_1, \sigma_1^2)$

$\vec{\theta}_2 = (\beta_1, \beta_2, \cdots, \beta_p, \nu_2, \sigma_2^2)$

Then $P_{\theta_1} = P_{\theta_2} \iff \left(\frac{1}{\sqrt{2\pi}\sigma_1}\right)^p = \left(\frac{1}{\sqrt{2\pi}\sigma_2}\right)^p \Rightarrow \sigma_1^2 = \sigma_2^2$

Also we get $\sum_{i=1}^{p}(X_i - \nu_1 - \alpha_i)^2 = \sum_{i=1}^{p}(X_i - \nu_2 - \beta_i)^2$

And we have the condition $\sum_{i=1}^{p} \alpha_i = \sum_{i=1}^{p} \beta_i = 0$

For any i, $V_1 + \alpha_i = V_2 + \beta_i \Rightarrow V_1 - V_2 = \beta_i - \alpha_i$

Then $p(V_1 - V_2) = \sum_{i=1}^{p} \beta_i - \sum_{i=1}^{p} \alpha_i = 0 \Rightarrow V_1 = V_2 \Rightarrow \beta_i = \alpha_i$

Thus, the parametrization is identifiable

3. $Y - X \sim N(\mu_2 - \mu_1, 2\sigma^2)$. So the pdf of $Y - X$ is

$$f_\theta(x) = \frac{1}{\sqrt{2\pi} \cdot 2\sigma^2} e^{-\frac{1}{4\sigma^4}(x - (\mu_2 - \mu_1))^2}.$$

Here we have a counterexample. Let $\theta_1 = (2,3)$, $\theta_2 = (3,4)$.
Then $Y - X \sim N(1, 2\sigma^2)$ but $\theta_1 \neq \theta_2$, so the parametrization is not identifiable.

4. Same as question 1, we can see that the parametrization is not identifiable.

5. $P_\theta = \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^{p+b} \cdot e^{\frac{-1}{2\sigma^2} \sum_{ij}^{p+b} (x_{ij} - \mu_{ij})^2}$, Let $\vec{\theta_1} = (\alpha_1^{(1)}, \dots, \lambda_1^{(1)}, \dots, V^{(1)}, \sigma_1^2)$,
$\vec{\theta_2} = (\alpha_1^{(2)}, \dots, \lambda_1^{(2)}, \dots, V^{(2)}, \sigma_2^2)$.

If $P_{\theta_1} = P_{\theta_2}$, Then $\sigma_1^2 = \sigma_2^2$, $\mu_{ij}^{(1)} = \mu_{ij}^{(2)} \Leftrightarrow V^{(1)} + \alpha_i^{(1)} + \lambda_j^{(1)} = V^{(2)} + \alpha_i^{(2)} + \lambda_j^{(2)}$

$V^{(1)} - V^{(2)} = \alpha_i^{(2)} - \alpha_i^{(1)} + \lambda_j^{(2)} - \lambda_j^{(1)}$, so $(p+b) \cdot (V^{(1)} - V^{(2)}) = b\left(\sum_{i=1}^{p} \alpha_i^{(2)} - \sum_{i=1}^{p} \alpha_i^{(1)}\right)$
$+ p\left(\sum_{j=1}^{b} \lambda_j^{(2)} - \sum_{j=1}^{b} \lambda_j^{(1)}\right) = 0$, which means $V^{(1)} = V^{(2)}$.

Then for any i, j $\alpha_i^{(1)} + \lambda_j^{(1)} = \alpha_i^{(2)} + \lambda_j^{(2)}$, Then $\alpha_i^{(1)} + \sum_{j=1}^{b} \lambda_j^{(1)} = \alpha_i^{(2)} + \sum_{j=1}^{b} \lambda_j^{(2)}$
$\Rightarrow \alpha_i^{(1)} = \alpha_i^{(2)}$, $\Rightarrow \lambda_j^{(1)} = \lambda_j^{(2)} \Rightarrow \theta_1 = \theta_2$ Thus the parametrization is identifiable.

$\square$

Problem B.0.2. Let $X_1, \ldots, X_n$ be a random sample from a population with density

$$f(x;\theta) = \left(\frac{x}{\theta^2}\right) \exp\left[-\frac{x^2}{2\theta^2}\right], \qquad x > 0, \qquad \theta > 0.$$

(a) Find the natural sufficient statistic. (b) Compute the expectation and variance of the natural sufficient statistic, in terms of the specified parameter θ and separately in terms of the canonical parameter η.

(a). Let $\frac{1}{\theta^2} = \eta$, we can write, $f(x,\eta) = \eta \cdot x \cdot \exp\left[\eta \cdot \left(-\frac{x^2}{2}\right)\right]$

$= x \cdot \exp\left[\eta \cdot \left(-\frac{x^2}{2}\right) + \log \eta\right] = h(x) \exp \cdot \left[\eta\, T(x) - A(\eta)\right]$, $|A(\eta)| < \infty$

Then the natural sufficient statistic $T(x) = -\frac{x^2}{2}$ $\quad A(\eta) = -\log \eta$

(b). $E[T(x)] = \int_0^\infty -\frac{x^3}{2\theta^2} \exp\left[-\frac{x^2}{2\theta^2}\right] dx = -\int_0^\infty x \cdot \exp\left[-\frac{x^2}{2\theta^2}\right] dx$

$= -\theta^2 = -\frac{1}{\eta} = A'(\eta)$

$E[T(x)^2] = \int_0^\infty \frac{x^5}{4\theta^2} \exp\left[-\frac{x^2}{2\theta^2}\right] = -\int_0^\infty x^3 \exp\left[-\frac{x^2}{2\theta^2}\right] dx$

$= \int_0^\infty 2\theta^2 x \cdot \exp\left[-\frac{x^2}{2\theta^2}\right] = 2\theta^4$

Then $Var[T(x)] = E[T(x)^2] - (E[T(x)])^2 = \theta^4 = \frac{1}{\eta^2} = A'(\eta)$.

$\square$

Problem B.0.4. Consider a random sample $X_1, \ldots, X_n$ from Multinomial$(1; \boldsymbol{\theta})$ where $\boldsymbol{\theta} \in \mathcal{S}_{k-1}$ is a vector of strictly positive probabilities summing to one. Here, $\mathcal{S}_{k-1}$ denotes the unit simplex. Show that the multinomial family is rank $k-1$. (Use this opportunity to (re)familiarize yourself with the Multinomial distribution, a generalization of the Binomial distribution.)

B.0.4, the pmf of X_i is $f(x) = P(m_1, m_2, \cdots, m_k \mid \vec{\theta})$

$$= \frac{1}{\prod_{i=1}^{k} m_i!} \cdot \prod_{i=1}^{k} \theta_i^{m_i} = \prod_{i=1}^{k} \theta_i^{m_i} = \exp\left(\sum_{i=1}^{k} m_i \log \theta_i\right)$$

supposing $\vec{\theta} = (\theta_1, \theta_2, \cdots, \theta_k)$ and $\sum_{i=1}^{k} \theta_i = 1$

Let $X = (m_1, \cdots, m_k)$ $\quad \eta = (\log \theta_1 \cdots \log \theta_k)$.

Then $f(X:\eta) = \exp(\eta \cdot X)$. where η satisfying $\sum_{i=1}^{k} \theta_i = 1$,

So we can let $\theta_k = 1 - \sum_{i=1}^{k-1} \theta_i$, so $\eta = (\log \theta_1, \cdots, \log \theta_{k-1}, \log(1 - \sum_{i=1}^{k-1} \theta_i))$

which means the rank of η is $k-1$, Thus the multinomial family is rank $k-1$.

$\square$

Problem B.0.5. Recall that the density function for a exponential random variable with parameter $\lambda > 0$ is given by $f_X(x) = \lambda \exp(-\lambda x)$ for $x \geq 0$. Prove that if $X_1 \sim \text{Exponential}(\lambda_1)$ and $X_2 \sim \text{Exponential}(\lambda_2)$ are independent, then $X = \min\{X_1, X_2\}$ follows Exponential(λ) with $\lambda = \lambda_1 + \lambda_2$.

B.0.5, $X_1 \sim$ Exponential (λ_1), $X_2 \sim$ Exponential (λ_2).

So we get that the cdf of X_1 is

$F(x;\lambda_1) = 1 - e^{-\lambda_1 x}$, Then we calculate the cdf of X

$$F(x;\lambda_1,\lambda_2) = P(t \leq x) = 1 - P(\min\{X_1, X_2\} > x)$$

$$= 1 - (P(X_1 > x) \times P(X_2 > x)) = 1 - (e^{-\lambda_1 x} \cdot e^{-\lambda_2 x})$$

$$= 1 - e^{-(\lambda_1+\lambda_2)x} = 1 - e^{-\lambda x}$$

Then the pdf of X is $\lambda \exp(-\lambda x)$. Thus X follows Exponential (λ).

$\square$