MTH 375
Fall 2022
Hw 3 key

#1.
4.2.2. Let $X_1, \ldots, X_n$ be a sample of iid Gamma(θ, 1) random variables with $\theta \in (0, \infty)$.
a) Determine the likelihood function $L(\theta)$.
$L(\theta) = \dfrac{1}{\Gamma(\theta)^n}(x_1 x_2 \cdots x_n)^{\theta-1} e^{-(x_1+\cdots+x_n)}$, $x_i > 0$.

b) Use the Fisher-Neyman factorization theorem to determine a sufficient statistic S for θ.
In Fisher-Neyman notation, $g(S, \theta) = \dfrac{1}{\Gamma(\theta)^n}(x_1 x_2 \cdots x_n)^{\theta-1}$ and $h(x_1, \ldots, x_n) = e^{-(x_1+\cdots+x_n)}$.
A sufficient statistic for θ is $S = X_1 X_2 \cdots X_n$.

#2. 4.2.4 Let $X_1, \cdots X_n$ be a sample of iid Beta(4,θ) random variables with $\theta \in (0, \infty)$.
a) Determine the likelihood function $L(\theta)$.
$L(\theta) = \left(\frac{\Gamma(4+\theta)}{\Gamma(4)\Gamma(\theta)}\right)^n (x_1 x_2 \cdots x_n)^3 \cdot ((1-x_1)(1-x_2)\cdots(1-x_n))^{\theta-1}$

b) In Fisher-Neyman notation, $g(S, \theta) = \left(\frac{\Gamma(4+\theta)}{\Gamma(4)\Gamma(\theta)}\right)^n ((1-x_1)(1-x_2)\cdots(1-x_n))^{\theta-1}$ and $h(x_1, \ldots, x_n) = (x_1 x_2 \cdots x_n)^3$. A sufficient statistic is $S = (1-X_1)(1-X_2)\cdots(1-X_n)$.

#3. 4.2.8 Let $X_1, \ldots, X_n$ be a sample of iid Beta(θ_1, θ_2) random variables with $\theta \in \mathbf{R}^+ \times \mathbf{R}^+$. Use the Fisher-Neyman factorization theorem to determine a sufficient statistic S for $\vec{\theta}$.
$L(\theta) = \left(\frac{\Gamma(\theta_1+\theta_2)}{\Gamma(\theta_1)\Gamma(\theta_2)}\right)^n (x_1 x_2 \cdots x_n)^{\theta_1-1} \cdot ((1-x_1)(1-x_2)\cdots(1-x_n))^{\theta_2-1}$.
Every term in $L(\theta)$ contains θ_1 or θ_2, so we need a pair of statistics for sufficiency, namely $(S_1, S_2) = (x_1 x_2 \cdots x_n, (1-x_1)(1-x_2)\cdots(1-x_n))$.

#4. 4.2.10. Let $X_1, , X_n$ be a sample of iid random variables with pdf $f(x) = \theta x^{\theta-1}$ for $x \in (0, 1)$ and $\theta \in (0, \infty)$. Use the Fisher-Neyman factorization theorem to determine a sufficient statistic for θ.
$L(\theta) = \theta^n (x_1 x_2 \cdots x_n)^{\theta-1}$. Here $g(S, \theta) = L(\theta)$ and $h(x_1, \ldots, x_n) = 1$. A sufficient statistic is $S = X_1 X_2 \cdots X_n$.

#5. Consider the family of distributions with pmf $p_X(x) = \begin{cases} p & \text{if } x = -1 \\ 2p & \text{if } x = 0 \\ 1-3p & \text{if } x = 1 \end{cases}$.

Here p is an unknown parameter, and $0 \le p \le 1/3$.

Let $X_1, X_2, \ldots, X_n$ be iid with common pmf a member of this family. Consider the statistics A= the number of i with $X_i = -1$, B= the number of i with $X_i = 0$, C= the number of i with $X_i = 1$.

(i) Write down the joint pmf of $X_1, X_2, \cdots, X_n$. This is most easily done using the statistics A, B, C. $p_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = p^A(2p)^B(1-3p)^C$.

(ii) Use the Fischer-Neyman Lemma to show that C is a sufficient statistic for p.

Since $A+B+C = n$, we can rewrite the joint pdf in the form
$p_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = p^{A+B}2^B(1-3p)^C = p^{n-C}2^B(1-3p)^C = p^{n-C}(1-3p)^C \cdot 2^B$.

In Fisher-Neyman notation, $g(C;p) = p^{n-C}(1-3p)^C$ and $h(x_1,\ldots,x_n) = 2^B$; C is a sufficient statistic for p.

(iii) Does it appear that A is also sufficient for p? Explain why or why not.

A is not sufficient for p. The factorization trick of (iii) that works for C does not work for A.

#6. Let X_1, X_2, X_3 be iid, binomial$(1,p)$. Let $T = X_1 + X_2 + 2X_3$. The purpose of this problem is to determine whether T is a sufficient statistic for p. Recall that the definition says that T is sufficient for p if for all $p \in [0,1]$

$$p_{X_1,X_2,X_3|T}(x_1,x_2,x_3 \mid T = x_1+x_2+2x_3) \quad \text{does not depend on } p.$$

Let's examine one particular instance of this definition.

Compute $p_{X_1,X_2,X_3|T}(0,0,1 \mid T=2)$. Does it depend on p? Is T a sufficient statistic for p?

Solution: There are only two triples for which $T = 2$, namely $(X_1,X_2,X_3) = (1,1,0)$ and $(0,0,1)$. Therefore $P(T=2) = p_{X_1,X_2,X_3}(1,1,0) + p_{X_1,X_2,X_3}(0,0,1) = p^2(1-p) + (1-p)^2p = p(1-p)\cdot(p+1-p) = p(1-p)$. Therefore

$$p_{X_1,X_2,X_3|T}(0,0,1 \mid T=2) = \frac{p_{X_1,X_2,X_3|T}(0,0,1)}{P(T=2)} = \frac{(1-p)^2p}{p(1-p)} = 1-p,$$

which certainly depends on p. We conclude that T is not a sufficient statistic for p.