541A Final Solutions[1]

1. Question 1

Let X be a random variable uniformly distributed in $[0,1]$.
(That is, X has PDF $f_X(x) = 1$ when $x \in [0,1]$, and $f_X(x) = 0$ when $x \notin [0,1]$.)
Let Y be a random variable uniformly distributed in $[0,1]$.
Assume that X and Y are independent.

- Compute $\mathbf{P}(X > 3/4)$.
- Compute $\mathbf{E}X$.
- Compute $\mathbf{P}(X + Y \leq 1/2)$.

Solution. Since X has PDF $f_X = 1_{[0,1]}$, we have $\mathbf{P}(X > 3/4) = \int_{3/4}^{1} dx = 1/4$ and $\mathbf{E}X = \int_0^1 x dx = 1/2$. Since X and Y are independent, they have joint PDF $f_{X,Y} = 1_{[0,1]^2}$, so that

$$\begin{aligned}\mathbf{P}(X+Y \leq 1) &= \int_{\{(x,y)\in\mathbf{R}^2\colon x\geq 0, y\geq 0, x+y\leq 1\}} = \int_{x=0}^{x=1/2}\int_{y=0}^{y=1/2-x} dydx \\ &= \int_{x=0}^{x=1/2}(1/2-x)dx = ((1/2)x - x^2/2)_{x=0}^{x=1/2} \\ &= (1/2)^2 - (1/2)^3 = 1/4 - 1/8 = 1/8.\end{aligned}$$

2. Question 2

Let $X_1, X_2, \ldots$ be real-valued random variables. Let X be a real-valued random variable with finite second moment. Assume that

$$\lim_{n\to\infty} \mathbf{E}\left|X_n - X\right|^2 = 0.$$

- Prove that $X_1, X_2, \ldots$ converges in probability to X.
- Does $X_1, X_2, \ldots$ converges in distribution to X? Justify your answer.

Solution. From Markov's inequality, we have, for any $\varepsilon > 0$,

$$\mathbf{P}(|X_n - X| > \varepsilon) \leq \frac{\mathbf{E}\left|X_n - X\right|^2}{\varepsilon^2}.$$

The right quantity converges to 0 as $n \to \infty$ by assumption. We therefore conclude that $X_1, X_2, \ldots$ converges in probability to X. Since convergence in probability implies convergence in distribution, $X_1, X_2, \ldots$ also converges in distribution to X.

3. Question 3

Let $n \geq 2$ be an integer. Let $X_1, \ldots, X_n$ be a random sample from the Gaussian distribution with mean $\mu \in \mathbf{R}$ and variance $\sigma^2 > 0$. That is, X_1 has PDF $\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $\forall$ $x \in \mathbf{R}$.

Let $\overline{X}_n := (X_1 + \cdots + X_n)/n$, and let $S_n := \sqrt{\frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2}$.

Show: $(n-1)S_n^2/\sigma^2$ is a chi-squared distributed random variable with $n-1$ degrees of freedom.

[1] May 3, 2023,

Hint: you can freely use the following fact:

$$nS_{n+1}^2 = (n-1)S_n^2 + \frac{n}{n+1}(X_{n+1} - \overline{X}_n)^2, \qquad \forall\, n \geq 2.$$

You can also freely use that S_n is independent of $\overline{X}_n$.

You can freely use that: a chi-squared random variable with k degrees of freedom has the same distribution as a sum squares of k i.i.d. standard Gaussians.

Solution. Let $\overline{X}_n := \frac{1}{n}\sum_{i=1}^n X_i$ and let $S_n^2 := \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$. In the case $n = 2$, we have $S_2^2 = \frac{1}{4}(X_1 - X_2)^2 + \frac{1}{4}(X_2 - X_1)^2 = \frac{1}{2}(X_1 - X_2)^2$. From Example 1.108 in the notes $\frac{1}{\sqrt{2}}(X_1 - X_2)$ is a mean zero Gaussian random variable with variance 1. So, S_2^2 is a chi-squared distributed random variable by the definition of a chi-squared random variable with one degree of freedom. That is, the third item of this proposition holds when $n = 2$. We now induct on n, using the hint.

Recall that S_n is independent of $\overline{X}_n$. Also, X_{n+1} is independent of S_n by Proposition 1.61 in the notes, since S_n is a function of $X_1, \ldots, X_n$, the latter being independent of X_{n+1}. In summary, S_n is independent of $(X_{n+1} - \overline{X}_n)^2$. By the inductive hypothesis, $(n-1)S_n^2$ is a chi-squared distributed random variable with $n-1$ degrees of freedom. From Example 1.108 in the notes, $X_{n+1} - \overline{X}_n$ is a Gaussian random variable with mean zero and variance $1 + 1/n$, so that $\sqrt{n/(n+1)}(X_{n+1} - \overline{X}_n)$ is a mean zero Gaussian with variance 1. The definition of a chi-squared random variable then implies that nS_{n+1} is a chi-squared random variable with n degrees of freedom, completing the inductive step.

4. Question 4

Let $X := (X_1, \ldots, X_n)$ be a random sample of size n from a binomial distribution with parameters n and p. Here n is a positive (known) integer and $0 < p < 1$ is unknown. (That is, $X_1, \ldots, X_n$ are i.i.d. and X_1 is a binomial random variable with parameters n and p, so that $\mathbf{P}(X_1 = k) = \binom{n}{k}p^k(1-p)^{n-k}$ for all integers $0 \leq k \leq n$.)

You can freely use that $\mathbf{E}X_1 = np$ and $\mathrm{Var}X_1 = np(1-p)$.

- Computer the Fisher information $I_X(p)$ for any $0 < p < 1$. (Consider n to be fixed.)
- Let Z be an unbiased estimator of p^2 (assume that Z is a function of $X_1, \ldots, X_n$). State the Cramér-Rao inequality for Z.
- Let W be an unbiased estimator of $1/p$ (assume that W is a function of $X_1, \ldots, X_n$). State the Cramér-Rao inequality for W.

Solution. Using that the information of independent random variables is the sum of the informations, using the alternate definition of Fisher information using the variance, and using that the variance is unchanged by adding a constant inside the variance,

$$
\begin{aligned}
I_X(p) = nI_{X_1}(p) &= n\mathrm{Var}_p\Big(\frac{d}{dp}\Big[\log\Big(\binom{n}{X_1}p^{X_1}(1-p)^{n-X_1}\Big)\Big]\Big)\\
&= n\mathrm{Var}_p\Big(\frac{d}{dp}\Big[\log\binom{n}{X_1} + X_1\log p + (n-X_1)\log(1-p)\Big]\Big)\\
&= n\mathrm{Var}_p\Big(\frac{d}{dp}\Big[X_1\log p + (n-X_1)\log(1-p)\Big]\Big)\\
&= n\mathrm{Var}_p\Big(\frac{1}{p}X_1 - \frac{1}{1-p}(n-X_1)\Big) = n\mathrm{Var}_p\Big(\Big[\frac{1}{p}+\frac{1}{1-p}\Big]X_1\Big)\\
&= n\Big[\frac{1}{p}+\frac{1}{1-p}\Big]^2\mathrm{Var}_p X_1 = n\Big[\frac{1}{p(1-p)}\Big]^2 np(1-p) = \frac{n^2}{p(1-p)}
\end{aligned}
$$

The Cramér-Rao inequality says, if $g(p) := \mathbf{E}_p Z$, then

$$\mathrm{Var}_p(Z) \geq \frac{|g'(p)|^2}{I_X(p)}.$$

If $g(p) = p^2$, then $g'(p) = 2p$, so we get

$$\mathrm{Var}_p(Z) \geq \frac{(2p)^2}{I_X(p)} = \frac{4p^3(1-p)}{n^2}.$$

If $g(p) = 1/p$, then $g'(p) = -p^{-2}$, so we get

$$\mathrm{Var}_p(Z) \geq \frac{p^{-4}}{I_X(p)} = p^{-3}\frac{1-p}{n^2}.$$

5. Question 5

Let $X_1, \ldots, X_n$ be i.i.d. random variables. Let $\theta > 0$ be an unknown parameter. Assume that X_1 is uniform on the interval $[0, \theta]$.

Denote $X := (X_1, \ldots, X_n)$.

- Is the Fisher Information $I_{X_1}(\theta)$ well-defined? Justify your answer. If $I_{X_1}(\theta)$ can be computed, simplify $I_{X_1}(\theta)$ to the best of your ability.
- Is the Fisher Information $I_X(\theta)$ well-defined? Justify your answer. If $I_X(\theta)$ can be computed, simplify $I_X(\theta)$ to the best of your ability.

Solution. $I_{X_1}(\theta)$ is not well-defined, since the region where the PDF of X_1 is nonzero is a function of θ. We could try to start with the definition of Fisher information with $f_\theta(x_1) = \theta^{-1}1_{[0,\theta]}$, so that $I_{X_1}(\theta) = \mathbf{E}_\theta(d/d\theta \log f_\theta(X_1))^2 = \mathbf{E}_\theta(d/d\theta\log\theta^{-1})^2 = \mathbf{E}_\theta(d/d\theta\log\theta^{-1})^2$

6. Question 6

Let $X_1, \ldots, X_n$ be i.i.d. random variables. Let $0 < p < 1$ be an unknown parameter. Assume that

$$\mathbf{P}(X_1 = k) = (1-p)^{k-1}p,$$

for all positive integers $k \geq 1$.

You can freely use the following facts: $\mathbf{E}X_1 = 1/p$, $\mathrm{Var}(X_1) = (1-p)/p^2$

- Find a statistic Z such that: Z is a method of moments estimator of p, and Z is a function of $X_1, \ldots, X_n$. Justify your answer.
- Find a statistic W such that W is an MLE for p, and W is a function of $X_1, \ldots, X_n$. (Make sure to justify that an MLE exists. Is an MLE unique in this case?) (Here MLE refers to a Maximum Likelihood Estimator.)

Solution. We have $\mathbf{E}X_1 = 1/p$, so that $p = 1/\mathbf{E}X_1$, so a Method of moments estimate of p is $Z = 1/((1/n)\sum_{i=1}^{n} X_i)$. The joint PMF of $X_1, \ldots, X_n$ is

$$f_p(x) = \prod_{i=1}^{n} (1-p)^{k_i - 1} p = p^n (1-p)^{-n} (1-p)^{\sum_{i=1}^{n} k_i}.$$

Viewed as a function of p, which we denote as $\ell(p)$, we have

$$\log \ell(p) = n \log p - n \log(1-p) + \log(1-p) \sum_{i=1}^{n} k_i.$$

Taking a derivative, we get

$$(\log \ell(p))' = n/p + n/(1-p) - (1/(1-p)) \sum_{i=1}^{n} k_i$$

Setting this equal to zero, we get

$$(1-p)n + pn - p \sum_{i=1}^{n} k_i = 0.$$

That is, $n - p\sum_{i=1}^{n} k_i = 0$, $p = n/\sum_{i=1}^{n} k_i$. So, $(\log \ell(p))' > 0$ for all $0 < p < n/\sum_{i=1}^{n} k_i$ and $(\log \ell(p))' < 0$ for all $p > n/\sum_{i=1}^{n} k_i$. (Since $k_1, \ldots, k_n \geq 1$, we have $n/\sum_{i=1}^{n} k_i \leq 1$ with equaliy only when $k_1 = \cdots = k_n = 1$.) So, the unique MLE exists, and it is equal to Z, so that $W = Z$.

7. Question 7

Let $X_1, \ldots, X_n$ be i.i.d. random variables. Let $\lambda > 0$ be an unknown parameter. Assume that

$$\mathbf{P}(X_1 = k) = (e^{-\lambda}) \frac{\lambda^k}{k!},$$

for all nonnegative integers $k \geq 0$.

You can freely use the following facts: $\mathbf{E}X_1 = \lambda$, $\mathrm{Var}(X_1) = \lambda$, $I_{X_1}(\lambda) = \frac{1}{\lambda}$.

You may assume that the MLE Y_n for λ exists and is unique, and it is given by:

$$Y_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

- Is Y_n unbiased for λ?
- Is the sequence of estimators $Y_1, Y_2, \ldots$ consistent? That is, does this sequence of random variables converge in probability to λ as $n \to \infty$? Justify your answer.
- What happens to the quantity $\mathrm{Var}(Y_n)$ as $n \to \infty$? More specifically, does this quantity asymptotically achieve the lower bound in the Cramér-Rao Inequality (for unbiased estimators of λ)? Justify your answer.

- Describe all unbiased estimators of λ that achieve equality in the Cramér-Rao inequality. (Assume such an estimator is a function of $X_1, \ldots, X_n$, for fixed n.)

Solution. Yes, Y_n is unbiased, since $\mathbf{E}X_1 = \lambda$, so $\mathbf{E}Y_n = \lambda$ as well. Consistency follows from the Weak Law of Large Numbers. The Cramér-Rao lower bound for unbiased estimators of λ would say that $\mathrm{Var}(Y_n) \geq 1/[nI_{X_1}(\lambda)] = \lambda/n$. And $\mathrm{Var}(Y_n) = \lambda/n$, so that Y_n does achieve equality in the Cramér-Rao inequality. Finally, we know that equality holds for an unbiased estimator Z of λ only when $Z - \mathbf{E}Z$ and $(d/d\lambda)\log f_\lambda(X)$ are multiples of each other, where

$$\frac{d}{d\lambda}\log f_\lambda(X) = \frac{d}{d\lambda}\log\prod_{i=1}^{n}(e^{-\lambda})\frac{\lambda^{X_i}}{X_i!} = \frac{d}{d\lambda}\Big(-\lambda n + \log\lambda\cdot\sum_{i=1}^{n}X_i\Big) = -n + \frac{1}{\lambda}\sum_{i=1}^{n}X_i\Big)$$

That is, there must exist a constant $c \in \mathbf{R}$ such that $Z - \lambda = c[-n + \frac{1}{\lambda}\sum_{i=1}^{n}X_i]$, i.e.

$$Z = \lambda - cn + \frac{c}{\lambda}\sum_{i=1}^{n}X_i.$$

Since Z is an estimator, if cannot have any factors of λ, and the only choice of c eliminating all of the λ factors is the choice $c = \lambda n$, so that $Z = Y_n$ is the unique unbiased estimator of λ achieving equality in the Cramér-Rao Inequality.

8. Question 8

Let $X_1, \ldots, X_n$ be i.i.d. random variables. Let $0 < p < 1$ be an unknown parameter. Assume that

$$\mathbf{P}(X_1 = k) = (1-p)^{k-1}p,$$

for all positive integers $k \geq 1$.

- Find a statistic Z such that: Z is complete for p, Z is sufficient for p, and Z is a function of $X_1, \ldots, X_n$. Justify your answer. (If you want to, you can use a result from the homework to do this part of the question.)
- Find a statistic W such that W is unbiased for p, and such that W is UMVU for p. Justify your answer.

(Hint: you can use that a sum of n independent geometric random variables has the same distribution as Y where $\mathbf{P}(Y = m) = \binom{m-1}{n-1}p^n(1-p)^{m-n}$ for all integers $m \geq n$.)

Solution. The joint PMF of $X_1, \ldots, X_n$ is

$$f_p(x) = \prod_{i=1}^{n}(1-p)^{k_i-1}p = p^n(1-p)^{-n}(1-p)^{\sum_{i=1}^{n}k_i}.$$

From the Factorization Theorem, we see that $\sum_{i=1}^{n}X_i$ is therefore a sufficient statistic. To see that this statistic is complete, we could either appeal to the homework (the result about complete statistics for exponential families), or we could argue as follows: if $\mathbf{E}_p f(Z) = 0$ for all $0 < p < 1$, then

$$0 = \mathbf{E}_p f(Z) = \sum_{m=n}^{\infty} f(m)\binom{m-1}{n-1}p^n(1-p)^{m-n}, \qquad \forall\, 0 < p < 1.$$

Multiplying by $p^{-n}(1-p)^n$, we get

$$0 = \mathbf{E}_p f(Z) = \sum_{m=n}^{\infty} f(m)\binom{m-1}{n-1}(1-p)^m, \qquad \forall\, 0<p<1.$$

The expression on the right is a power series in $1-p$ that converges for all $0<p<1$, and it is equal to zero, so all of its coefficients must be zero, i.e. $f(Z)=0$, so that Z is complete for p. Now, since $1_{X_1=1}$ satisfies $\mathbf{E}1_{X_1=1} = \mathbf{P}(X_1=1) = p$, the Lehmann-Scheffé Theorem implies that the following statistic is UMVU for p:

$$\mathbf{E}(1_{X_1=1}|\sum_{i=1}^{n} X_i).$$

To compute this conditional expectation, we perform the following computation:

$$\mathbf{E}(1_{X_1=1}|\sum_{i=1}^{n} X_i = t) = \mathbf{P}(X_1=1\,|\,\sum_{i=1}^{n} X_i = t) = \mathbf{P}(X_1=1, \sum_{i=1}^{n} X_i = t)/\mathbf{P}(\sum_{i=1}^{n} X_i = t)$$
$$= \mathbf{P}(X_1=1, \sum_{i=2}^{n} X_i = t-1)/\mathbf{P}(\sum_{i=1}^{n} X_i = t)$$
$$= p\frac{\binom{t-2}{n-2}p^{n-1}(1-p)^{m-n+1-1}}{\binom{t-1}{n-1}p^n(1-p)^{m-n}} = \frac{\binom{t-2}{n-2}}{\binom{t-1}{n-1}} = \frac{(t-2)!}{(n-2)!(n-t)!}\frac{(n-1)!(n-t)!}{(t-1)!} = \frac{n-1}{t-1}.$$

So, the UMVU is

$$\mathbf{E}(1_{X_1=1}|\sum_{i=1}^{n} X_i) = \frac{n-1}{-1+\sum_{i=1}^{n} X_i}$$

At least when $n>1$. (When $n=1$ this expression is zero, so it is not the UMVU.) In the case $n=1$, the statistic $1_{X_1=1}$ itself is UMVU for p, since it is unbiased and it is a function of the complete, sufficient statistic X_1.