408 Final Exam Solutions[1]

1. Question 1

(a) Let X be a random variable with $\mathbf{P}(X=-3)=1/3$ and $\mathbf{P}(X=3)=2/3$. Compute $\mathbf{E}X$ and $\mathbf{E}(X^2)$.

By definition of X, we have $\mathbf{E}X=-3(1/3)+3(2/3)=-1+2=1$ and $\mathbf{E}X^2=(-3)^2(1/3)+3^2(2/3)=9$.

(b) Let Y,Z be i.i.d. random variables. Assume that Y is uniformly distributed in $[0,1]$. Compute $\mathbf{P}(Y+Z\leq 1/2)$.

Since (Y,Z) is uniformly distributed in $[0,1]^2$, $\mathbf{P}(Y+Z\leq 1/2)$ is the area of a right triangle with edge lengths $1/2$, i.e. $\mathbf{P}(Y+Z\leq 1/2)=(1/2)^3=1/8$.

(c) State the Central Limit Theorem. Make sure to include **all** assumptions.

Let $X_1,\ldots,X_n$ be independent identically distributed random variables. Assume that $\mathbf{E}\,|X_1|<\infty$ and $0<\mathrm{Var}(X_1)<\infty$.

Let $\mu=\mathbf{E}X_1$ and let $\sigma=\sqrt{\mathrm{Var}(X_1)}$. Then for any $-\infty\leq a\leq\infty$,

$$\lim_{n\to\infty}\mathbf{P}\left(\frac{X_1+\cdots+X_n-\mu n}{\sigma\sqrt{n}}\leq a\right)=\int_{-\infty}^{a}e^{-t^2/2}\frac{dt}{\sqrt{2\pi}}.$$

2. Question 2

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

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

[1]December 13, 2023, © 2023 Steven Heilman, All Rights Reserved.

Setting this equal to zero, we get

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

That is, $n - p\sum_{i=1}^{n} k_i = 0$, $p = n/\sum_{i=1}^{n} k_i$. Evidently, $(\log \ell(p))'$ is increasing for p below this value and increasing for p above this value. We conclude that the unique MLE for p is

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

3. Question 3

Let $\theta \in \mathbf{R}$ be an unknown parameter. Consider the PDF

$$f_\theta(x) := \begin{cases} e^{-(x-\theta)}, & \text{if } x \geq \theta \\ 0, & \text{if } x < \theta. \end{cases}$$

Suppose $X_1, \ldots, X_n$ is a random sample of size n, such that X_i has PDF f_θ for all $1 \leq i \leq n$.

Show that $X_{(1)} = \min_{1\leq i\leq n} X_i$ is a sufficient statistic for θ. *Solution.* Let $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$. If it occurs that $\min_{1\leq i\leq n} x_i < \theta$, then some $1 \leq i \leq n$ satisfies $x_i < \theta$, so $f_\theta(x_i) = 0$ and the joint PDF $\prod_{i=1}^{n} f_\theta(x_i)$ is also zero. On the other hand, if it occurs that $\min_{1\leq i\leq n} x_i \geq \theta$, then all $1 \leq i \leq n$ satisfy $x_i \geq \theta$, and the joint PDF $\prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n} e^{-(x_i-\theta)}$. That is, we can write

$$\prod_{i=1}^{n} f_\theta(x_i) = 1_{\{\min_{1\leq i\leq n} x_i \geq \theta\}} \cdot \prod_{i=1}^{n} e^{-(x_i-\theta)} = e^{n\theta} 1_{\{\min_{1\leq i\leq n} x_i \geq \theta\}} \cdot \prod_{i=1}^{n} e^{-(x_i)}.$$

The factorization theorem then implies that $t(x) := \min_{1\leq i\leq n} x_i$ gives our sufficient statistic $Y = t(X_1, \ldots, X_n)$, since if we define $g_\theta(z) := e^{n\theta} 1_{\{z\geq\theta\}}$ and $h(x) := \prod_{i=1}^{n} e^{-(x_i)}$, then we have written the joint PDF as

$$\prod_{i=1}^{n} f_\theta(x_i) = g_\theta(t(x))h(x), \qquad \forall\, x \in \mathbf{R}^n, \quad \forall\, \theta \in \mathbf{R}.$$

4. Question 4

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.

5. Question 5

Suppose X is a binomial distributed random variable with parameters 2 and $\theta \in \{1/2, 3/4\}$. (So, X has the distribution of the number of heads that appears from flipping a coin twice, where θ is the probability that a heads appears in a single coin flip.)

We want to test the hypothesis H_0 that $\theta = 1/2$ versus the hypothesis H_1 that $\theta = 3/4$.

- Explicitly describe the rejection region C of the UMP (uniformly most powerful) test among all hypothesis tests with significance level at most $1/4$.
- Explicitly describe the rejection region C of the UMP (uniformly most powerful) test among all hypothesis tests with significance level at most $3/4$.
- Suppose we observe that $X = 2$. Report a p-value for this observation, for the UMP tests you found.

Solution. The Neyman-Pearson Lemma says that the UMP test for the class of tests with an upper bound on the significance level must be a likelihood ratio test. There are only three values that X can take, so we examine the likelihood ratios explicitly:

$$\frac{f_{3/4}(0)}{f_{1/2}(0)} = \frac{(1-3/4)^2}{(1-1/2)^2} = \frac{1}{4}, \qquad \frac{f_{3/4}(1)}{f_{1/2}(1)} = \frac{2(1-3/4)(3/4)}{2(1-1/2)(1/2)} = \frac{3}{4}, \qquad \frac{f_{3/4}(2)}{f_{1/2}(2)} = \frac{(3/4)^2}{(1/2)^2} = \frac{9}{4}.$$

We then get different likelihood ratio tests according to the choice of $k > 0$.

- If $3/4 < k \leq 9/4$, then H_0 is rejected if and only if $X = 2$, and this test is the unique UMP for tests with significance level at most $\mathbf{P}_{1/2}(X = 2) = 1/4$.
- If $1/4 < k \leq 3/4$, then H_0 is rejected if and only if $X = 1$ or 2, and this test is the unique UMP for tests with significance level at most $\mathbf{P}_{1/2}(X \in \{1, 2\}) = 3/4$.
- If $0 < k \leq 1/4$, then H_0 is always rejected, and this test is the unique UMP for tests with significance level at most $\mathbf{P}_{1/2}(X \in \{1, 2, 3\}) = 1$.
- If $k > 9/4$, then H_0 is never rejected, and this test is the unique UMP for tests with significance level at most $\mathbf{P}_{1/2}(X \in \emptyset) = 0$.

In the case $k = 2$, we get (using the table of values of $f_{3/4}(X)/f_{1/2}(X)$),

$$p(2) = \mathbf{P}_{1/2}\Big(\frac{f_{3/4}(X)}{f_{1/2}(X)} \geq \frac{f_{3/4}(2)}{f_{1/2}(2)}\Big) = \mathbf{P}_{1/2}\Big(\frac{f_{3/4}(X)}{f_{1/2}(X)} \geq \frac{9}{4}\Big) = \mathbf{P}_{1/2}(X = 2) = (1/2)^2 = 1/4.$$

6. Question 6

Suppose you flip a coin 1000 times, resulting in 600 heads and 400 tails. Is it reasonable to conclude that the coin is fair (i.e. it has one half probability of heads and one half probability of tails)? Justify your answer. (Your answer should compute a p-value for a sensible hypothesis test.)

Solution. Let X be the number of heads that were flipped. Under the null hypothesis, X has a binomial distribution with parameters $n = 1000$ and $\theta = 1/2$. Consider the hypothesis test that rejects when $|X - 500| \geq 100$. From the Chebyshev Inequality,

$$\mathbf{P}_{1/2}(|X - 500| \geq 100) \leq \frac{\mathrm{Var}(X)}{100^2} = \frac{1000(1/4)}{10^4} = \frac{1}{40} \approx .025.$$

So, the p-value for the observation that $X = 600$ is

$$p(600) = \mathbf{P}_{1/2}(|X - 500| \geq |600 - 500|) \leq .025.$$

Since this p-value is small, we are confident in rejecting the null hypothesis. That is, the coin is probably not fair.

7. Question 7

Suppose X_1, X_2, X_3 are i.i.d. Gaussian random variables with variance one and unknown mean $\mu \in \mathbf{R}$. (A Gaussian with mean μ and variance $\sigma^2 > 0$ has PDF $\frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/[2\sigma^2]}$.) Suppose you find that $X_1 = 1$ and $X_2 = 3$ and $X_3 = 0$.

Explicitly construct a confidence interval of the form $[a, b]$ for μ, so that

$$\mathbf{P}(a \leq \mu \leq b) = \int_{-1}^{1} \frac{1}{\sqrt{3}\sqrt{2\pi}} e^{-y^2/6} dy.$$

(Hint: What is the PDF of $X_1 + X_2 + X_3 - 3\mu$?)

Solution. Since X_1, X_2, X_3 are i.i.d. Gaussians, $X_1 + X_2 + X_3 - 3\mu$ is a mean zero Gaussian with variance 3. That is,

$$\mathbf{P}(a \leq X_1 + X_2 + X_3 - 3\mu \leq b) = \int_a^b \frac{1}{\sqrt{3}\sqrt{2\pi}} e^{-y^2/6} dy.$$

Rearranging, we get

$$\begin{aligned}\int_a^b \frac{1}{\sqrt{3}\sqrt{2\pi}} e^{-y^2/9} dy &= \mathbf{P}(a - (X_1 + X_2 + X_3) \leq -3\mu \leq b - (X_1 + X_2 + X_3)) \\ &= \mathbf{P}([-a + X_1 + X_2 + X_3]/3 \geq \mu \geq [-b + X_1 + X_2 + X_3]/3)\end{aligned}$$

Choosing $a = -1$ and $b = 1$ we get

$$\mathbf{P}([1 + X_1 + X_2 + X_3]/3 \geq \mu \geq [-1 + X_1 + X_2 + X_3]/3) = \int_{-1}^{1} \frac{1}{\sqrt{3}\sqrt{2\pi}} e^{-y^2/6} dy$$

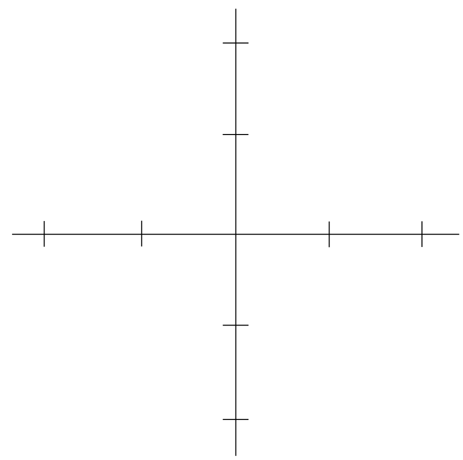
We also have $[1 + X_1 + X_2 + X_3]/3 = 5/3$ and $[-1 + X_1 + X_2 + X_3]/3 = 1$ That is, we choose $[a, b] = [1, 5/3]$.

8. Question 8

Suppose you are given the following three data points in (x, y) coordinates:

$$(x_1, y_1) = (-1, 0), \qquad (x_2, y_2) = (0, 0), \qquad (x_3, y_3) = (0, 2).$$

- Find the line of the form $y = mx + b$ that best fits these three points. That is, find $m, b \in \mathbf{R}$ that minimizes the quantity.
$$h(m, b) := \frac{1}{2}\sum_{i=1}^{3}\Big(y_i - (mx_i + b)\Big)^2.$$
- Make sure to prove that the minimal m, b that you find actually minimizes $h(m, b)$.
- Finally, plot the points (x_1, y_1), (x_2, y_2), (x_3, y_3) along with the line $y = mx + b$ that best fits the points.



Solution. We have $h(m, b) = (1/2)[(-m + b)^2 + b^2 + (2 - b)^2]$. Then $h_m(m, b) = m - b$ and $h_b(m, b) = b - m + b + b - 2 = 3b - m - 2$. Solving for $h_m = h_b = 0$, we get $m = b$ and $0 = 3b - m - 2 = 2b - 2$, so that $b = m = 1$. So, the parameters $b = 1$, $m = 1$ are the only critical point of h. Since h is strictly convex, its critical point must be a global minimum.

The "best-fit" line is then

$$y = x + 1.$$

Note that $y(-1) = 0$, $y(0) = 1$. So, the line intersects $(-1, 0)$, it lies above $(0, 0)$ and it lies below $(0, 2)$.