Time: 08.00-13.00. Total sum: 40p. Grades 3, 4 and 5 require 18p, 25p and 32p, respectively. Note:

- Motivate solutions well but no more than necessary.
- Just write on one side of the paper.
- Start new question on new page, and label the question number clearly next to your solution.
- Do not write with red-ink pen.
- Write the solutions in increasing order of question numbers.
- Hints and other information provided in a problem may also be useful for subsequent problems.

Permitted aids: Any books, notes, and pocket calculator.

Ex. 1 — *5p* – Eight earthquakes with the largest magnitude (in Richter scale) that hit Christchurch, NZ on 22nd February 2011 are:

$$5.8450, 4.6800, 5.9100, 5.0110, 4.7550, 4.8350, 4.6310, 4.6400$$

(a)— *1p* – Report the sample mean.

(b)— *1p* – Report the sample variance and sample standard deviation.

(c)— *1p* – Report the five order statistics, as a point in $\mathbb{R}^5$, from minimum to maximum.

(d)— *2p* – Sketch the Empirical Distribution Function showing discontinuities in the function and clearly labeling the axes.

Ex. 2 — *5p* – Suppose you plan to obtain an independent and identically distributed (IID) sequence of n measurements from an instrument. This instrument has been calibrated so that the distribution of measurements made with it have population variance of $1/4$. Possibly useful fact: If $1-\alpha = 0.95$ and $Z \sim N(0,1)$ is the standard Normal RV, then from the Z Table, $z_{\alpha/2} = 1.96$, where, $P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1-\alpha$. If $X \sim Poisson(\lambda)$ then $V(X) = E(X) = \lambda$]

(a)— *4p* – Your boss wants you to make a point estimate of the unknown population mean from an IID sample of size n. He also insists that the tolerance for error has to be $1/10$ and the probability of meeting this tolerance should be just above 95%. Use the Central Limit Theorem to find how large should n be to meet the specifications of your boss. In summary,

$$X_1, X_2, \ldots, X_n \overset{IID}{\sim} X_1, \quad V(X_1) = 1/4$$

Find n such that $P(|\bar{X}_n - E(X_1)| < 1/10) = 0.95$.

(b)— *1p* – Making the further assumption that the IID samples are Poisson distributed random variables, find the Method of Moments Estimate for the mean parameter λ of the $Poisson(\lambda)$ RV X_i whose probability mass function is given by:

$$f(x; \lambda) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k \in \{0, 1, 2, \ldots\}, \quad \lambda > 0$$

Ex. 3 — *5 p* – Assume that an independent and identically distributed sample, $X_1, X_2, \ldots, X_n$ is drawn from the distribution of X with PDF $f(x; \theta)$:

$$f(x; \theta) = \begin{cases} \theta x^{\theta - 1} & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} .$$

for a fixed and unknown parameter $\theta \in (0, \infty)$ and derive the maximum likelihood estimate of θ.

Hint: you only need to do Steps 1–5 to find the MLE: (Step 1:) find $\ell(\theta)$, the log-likelihood as a function of the parameter θ, (Step 2:) find $\frac{d}{d\theta}\ell(\theta)$, the first derivative of $\ell(\theta)$ with respect to θ, (Step 3:) solve the equation $\frac{d}{d\theta}\ell(\theta) = 0$ for θ and set this solution equal to $\widehat{\theta}_n$, (Step 4:) find $\frac{d^2}{d\theta^2}\ell(\theta)$, the second derivative of $\ell(\theta)$ with respect to θ and finally (Step 5:) $\widehat{\theta}_n$ is the MLE if $\frac{d^2}{d\theta^2}\ell(\theta) < 0$.

Ex. 4 — *15p* – Let $X_1, X_2, \ldots, X_n \overset{IID}{\sim} Normal(\mu, \sigma^2)$. Suppose that μ is known and σ is unknown. The parameter of interest is $\psi = \log(\sigma)$. Answer the following:

(1)— *2p* – Find the log-likelihood function $\ell(\sigma)$

(2)— *2p* – Find its derivative with respect to the unknown parameter σ

(3)— *2p* – Set the derivative equal to 0 and solve for σ

(4)— *2p* – Find the estimated standard error $\widehat{se}_n$ for the estimator of σ via Fisher Information

(5)— *2p* – Derive the estimated standard error of $\psi = \log(\sigma)$ via the Delta method

(6)— *2p* – Obtain the 95% confidence interval for ψ

(7)— *1p* – Suppose you observed $n = 110$ samples and a sample standard deviation of 12.4, Will you reject or fail to reject the null hypothesis $H_0 : \psi = 10.0$ versus $H_1 : \psi \neq 10.0$ using a size $\alpha = 0.05$ Wald test?

(8)— *1p* – Obtain the 95% confidence interval for σ based on $n = 110$ samples and a sample standard deviation of 12.4

(9)— *1p* – Will you reject or fail to reject the null hypothesis $H_0 : \sigma = 10.0$ versus $H_1 : \sigma \neq 10.0$ using a size $\alpha = 0.05$ Wald test?

Ex. 5 — *5p* – Let $X_1, X_2, \ldots, X_n \overset{IID}{\sim} Poisson(\lambda)$ where $\lambda = E(X_i) > 0$ is unknown. Show that the sample mean $\bar{X}_n = n^{-1} \sum_{i=1}^{n} X_i$ is a sufficient statistic.

Ex. 6 — *5p* – Suppose $X_1, X_2, \ldots, X_m \overset{IID}{\sim} F$ where F is any distribution function of a real-valued random variable. Recall that a permutation can be defined as a bijection from a set onto itself, i.e., $\pi : \{1, 2, \ldots, m\} \to \{1, 2, \ldots, m\}$.

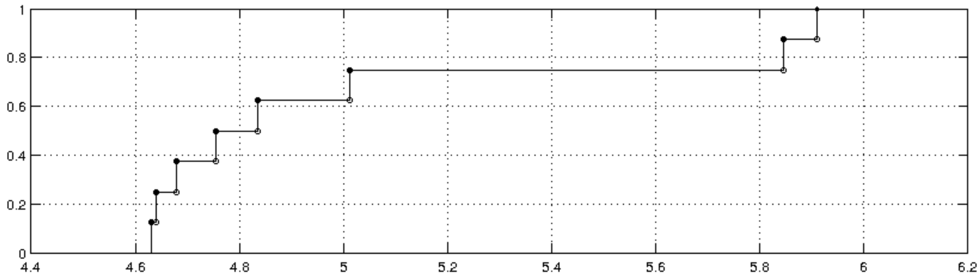
(a)– *2p* — Show that $\widehat{F}_n(x) := n^{-1} \sum_{i=1}^{n} \mathbf{1}(X_i \leq x)$ is an unbiased estimator of $F(x)$, for each $x \in \mathbb{R}$, where,

$$\mathbf{1}(X_i \leq x) = \begin{cases} 1 \text{ if } X_i \leq x \\ 0 \text{ if } X_i > x \end{cases}$$

(b)– *2p* — What is the Variance of the estimator $\widehat{F}_n(x)$ of $F(x)$, for each $x \in \mathbb{R}$.

(c)– *1p* — Why is $P(X_1 = x_1, X_2 = x_2, \ldots, X_m = x_m) = P(X_1 = x_{\pi(1)}, X_2 = x_{\pi(2)}, \ldots, X_m = x_{\pi(m)}) = 1/m!$? Explain your answer step by step.

Lycka till!



Figur 1: Empirical Distribution Function

Answer (Ex. 1) —

(a)sample mean = 5.0384

(b)sample variance and sample standard deviation are: $0.2837, 0.5326$, respectively.

(c)order statistics is:

$$(4.6310, 4.6400, 4.6800, 4.7550, 4.8350, 5.0110, 5.8450, 5.9100),$$

(d)Empirical Distribution Function is given in Fig. 1.

Answer (Ex. 2) —

(a)Using the CLT and the given fact:

$$P(|\bar{X}_n - E(X_1)| < 0.10) = 0.95$$
$$P(-0.10 < \bar{X}_n - E(X_1) < 0.10) = 0.95$$

(1p)

$$P\left(\frac{-0.10}{\sqrt{V(X_1)/n}} < \frac{\bar{X}_n - E(X_1)}{\sqrt{V(X_1)/n}} < \frac{0.10}{\sqrt{V(X_1)/n}}\right) = 0.95$$
$$P\left(\frac{-0.10}{\sqrt{1/4n}} < Z < \frac{0.10}{\sqrt{1/4n}}\right) = 0.95$$

(1p)

We want $1-\alpha = 0.95$, and from the standard Normal Table we know that the corresponding $z_{\alpha/2} = 1.96$. So, $\frac{0.10}{\sqrt{1/4n}} = z_{\alpha/2} = 1.96$ (1p)
and therefore we can get the right sample size n as follows:

$$\begin{aligned} n &= \left((\sqrt{1/4} \times 1.96)/(1/10) \right)^2 \\ &= (((1/2) \times 1.96)/(1/10))^2 = (0.98 \times 10)^2 = 9.8^2 = 96.04 \end{aligned}$$

Finally, by rounding 96.04 up to the next largest integer we need $n = 97$ measurements to meet the specifications of your boss (at least up to the approximation provided by the CLT). (1p)

(b)

$$E(X_i; \lambda) = \lambda = \bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

The Method of Moment Estimator of λ is $\bar{X}_n$. (1p)

Answer (Ex. 3) — Step 1: If $x_i \in (0,1)$ for each $i \in \{1, 2, \ldots, n\}$, i.e. when each data point lies inside the open interval $(0,1)$, the log-likelihood is

$$\begin{aligned} \ell(\theta) = \sum_{i=1}^{n} \log(f_X(x_i;\theta)) &= \sum_{i=1}^{n} \left(\log\left(\theta x_i^{\theta-1}\right)\right) = \sum_{i=1}^{n} \left(\log(\theta) + \log\left(x_i^{\theta-1}\right)\right) \\ &= \sum_{i=1}^{n} (\log(\theta) + (\theta-1)(\log(x_i))) = \sum_{i=1}^{n} (\log(\theta) + \theta \log(x_i) - \log(x_i)) \end{aligned}$$

$$= \sum_{i=1}^{n} \log(\theta) + \sum_{i=1}^{n} \theta \log(x_i) - \sum_{i=1}^{n} \log(x_i) = n \log(\theta) + \theta \sum_{i=1}^{n} \log(x_i) - \sum_{i=1}^{n} \log(x_i)$$

Step 2:

$$\frac{d}{d\theta}(\ell(\theta)) = \frac{d}{d\theta}\left(n\log(\theta) + \theta \sum_{i=1}^{n} \log(x_i) - \sum_{i=1}^{n} \log(x_i) \right) = \frac{n}{\theta} + \sum_{i=1}^{n} \log(x_i) - 0 = \frac{n}{\theta} + \sum_{i=1}^{n} \log(x_i)$$

Step 3:

$$\frac{d}{d\theta}(\ell(\theta)) = 0 \iff \frac{n}{\theta} + \sum_{i=1}^{n} \log(x_i) = 0 \iff \frac{n}{\theta} = -\sum_{i=1}^{n} \log(x_i) \iff \theta = -\frac{n}{\sum_{i=1}^{n} \log(x_i)}$$

Let

$$\widehat{\theta}_n = -\frac{n}{\sum_{i=1}^n \log(x_i)} \quad .$$

Step 4:

$$\frac{d^2}{d\theta^2}\ell(\theta) = \frac{d}{d\theta}\left(\frac{d}{d\theta}(\ell(\theta))\right) = \frac{d}{d\theta}\left(\frac{n}{\theta} + \sum_{i=1}^n \log(x_i)\right) = \frac{d}{d\theta}\left(n\theta^{-1} + \sum_{i=1}^n \log(x_i)\right)$$
$$= -n\theta^{-2} + 0 = -\frac{n}{\theta^2}$$

Step 5: The problem states that $\theta > 0$. Since $\theta^2 > 0$ and $n \geq 1$, we have indeed checked that

$$\frac{d^2}{d\theta^2}\ell(\theta) = -\frac{n}{\theta^2} < 0$$

and therefore the MLE is indeed

$$\widehat{\theta}_n = \frac{-n}{\sum_{i=1}^n \log(x_i)} \quad .$$

Answer (Ex. 4) — (a)— 2p –

$$\ell(\sigma) := \log(L(\sigma)) := \log(L(x_1, x_2, \ldots, x_n; \sigma)) = \log\left(\prod_{i=1}^n f(x_i;\sigma)\right) = \sum_{i=1}^n \log(f(x_i;\sigma))$$
$$= \sum_{i=1}^n \log\left(\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(x_i-\mu)^2\right)\right)$$
$$= \sum_{i=1}^n \left(\log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) + \log\left(\exp\left(-\frac{1}{2\sigma^2}(x_i-\mu)^2\right)\right)\right)$$
$$= \sum_{i=1}^n \log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) + \sum_{i=1}^n\left(-\frac{1}{2\sigma^2}(x_i-\mu)^2\right) = n\log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) + \left(-\frac{1}{2\sigma^2}\right)\sum_{i=1}^n (x_i-\mu)^2$$
$$= n\left(\log\left(\frac{1}{\sqrt{2\pi}}\right) + \log\left(\frac{1}{\sigma}\right)\right) - \left(\frac{1}{2\sigma^2}\right)\sum_{i=1}^n (x_i-\mu)^2$$
$$= n\log\left(\sqrt{2\pi}^{-1}\right) + n\log\left(\sigma^{-1}\right) - \left(\frac{1}{2\sigma^2}\right)\sum_{i=1}^n (x_i-\mu)^2$$
$$= -n\log\left(\sqrt{2\pi}\right) - n\log(\sigma) - \left(\frac{1}{2\sigma^2}\right)\sum_{i=1}^n (x_i-\mu)^2$$

(b)— *2p* –

$$\begin{aligned}\frac{\partial}{\partial \sigma}\ell(\sigma) &:= \frac{\partial}{\partial \sigma}\left(-n\log\left(\sqrt{2\pi}\right) - n\log(\sigma) - \left(\frac{1}{2\sigma^2}\right)\sum_{i=1}^{n}(x_i-\mu)^2\right)\\ &= \frac{\partial}{\partial \sigma}\left(-n\log\left(\sqrt{2\pi}\right)\right) - \frac{\partial}{\partial \sigma}(n\log(\sigma)) - \frac{\partial}{\partial \sigma}\left(\left(\frac{1}{2\sigma^2}\right)\sum_{i=1}^{n}(x_i-\mu)^2\right)\\ &= 0 - n\frac{\partial}{\partial \sigma}(\log(\sigma)) - \left(\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right)\frac{\partial}{\partial \sigma}\left(\sigma^{-2}\right)\\ &= -n\sigma^{-1} - \left(\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right)\left(-2\sigma^{-3}\right) = -n\sigma^{-1} + \sigma^{-3}\sum_{i=1}^{n}(x_i-\mu)^2\end{aligned}$$

(c)— *2p* –

$$\begin{aligned}0 = -n\sigma^{-1} + \sigma^{-3}\sum_{i=1}^{n}(x_i-\mu)^2 \iff n\sigma^{-1} = \sigma^{-3}\sum_{i=1}^{n}(x_i-\mu)^2 &\iff n\sigma^{-1}\sigma^{+3} = \sum_{i=1}^{n}(x_i-\mu)^2\\ \iff n\sigma^{-1+3} = \sum_{i=1}^{n}(x_i-\mu)^2 &\iff n\sigma^{2} = \sum_{i=1}^{n}(x_i-\mu)^2\\ \iff \sigma^2 = \left(\sum_{i=1}^{n}(x_i-\mu)^2\right)/n &\iff \sigma = \sqrt{\sum_{i=1}^{n}(x_i-\mu)^2/n}\end{aligned}$$

So MLE $\widehat{\sigma_n} = \sqrt{\sum_{i=1}^{n}(x_i-\mu)^2/n}$.

(d)— *2p* – The Log-likelihood function of σ, based on one sample from the $Normal(\mu, \sigma^2)$ RV with known μ is,

$$\log f(x;\sigma) = \log\left(\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)\right) = -\log\left(\sqrt{2\pi}\right) - \log(\sigma) - \left(\frac{1}{2\sigma^2}\right)(x-\mu)^2$$

Therefore, in much the same way as in part (2) earlier,

$$\begin{aligned}\frac{\partial^2 \log f(x;\sigma)}{\partial^2 \sigma} &:= \frac{\partial}{\partial \sigma}\left(\frac{\partial}{\partial \sigma}\left(-\log\left(\sqrt{2\pi}\right) - \log(\sigma) - \left(\frac{1}{2\sigma^2}\right)(x-\mu)^2\right)\right)\\ &= \frac{\partial}{\partial \sigma}\left(-\sigma^{-1} + \sigma^{-3}(x-\mu)^2\right) = \sigma^{-2} - 3\sigma^{-4}(x-\mu)^2\end{aligned}$$

Now, we compute the Fisher Information of one sample as an expectation of the continuous RV X over $(-\infty, \infty)$ with density $f(x;\sigma)$,

$$\begin{aligned}
I_1(\sigma) &= -\int_{x\in(-\infty,\infty)} \left(\frac{\partial^2 \log f(x;\sigma)}{\partial^2 \lambda}\right) f(x;\sigma)\ dx = -\int_{-\infty}^{\infty} \left(\sigma^{-2} - 3\sigma^{-4}(x-\mu)^2\right) f(x;\sigma)\ dx \\
&= \int_{-\infty}^{\infty} -\sigma^{-2} f(x;\sigma)\ dx + \int_{-\infty}^{\infty} 3\sigma^{-4}(x-\mu)^2 f(x;\sigma)\ dx \\
&= -\sigma^{-2} \int_{-\infty}^{\infty} f(x;\sigma)\ dx + 3\sigma^{-4} \int_{-\infty}^{\infty} (x-\mu)^2 f(x;\sigma)\ dx \\
&= -\sigma^{-2} + 3\sigma^{-4}\sigma^2 \qquad \because \sigma^2 = V(X) = E(X - E(X))^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x;\sigma)\ dx \\
&= -\sigma^{-2} + 3\sigma^{-4+2} = -\sigma^{-2} + 3\sigma^{-2} = 2\sigma^{-2}
\end{aligned}$$

Therefore, the estimated standard error of the estimator of the unknown σ is

$$\frac{1}{\sqrt{I_n(\widehat{\sigma}_n)}} = \frac{1}{\sqrt{nI_1(\widehat{\sigma}_n)}} = \frac{1}{\sqrt{n2(\widehat{\sigma}_n)^{-2}}} = \frac{\widehat{\sigma}_n}{\sqrt{2n}} .$$

(e)— *2p* –

$$\widehat{\mathsf{se}}_n(\widehat{\Psi}_n) = |g'(\sigma)| \widehat{\mathsf{se}}_n(\widehat{\sigma}_n) = \left|\frac{\partial}{\partial\sigma}\log(\sigma)\right| \frac{\sigma}{\sqrt{2n}} = \frac{1}{\sigma}\frac{\sigma}{\sqrt{2n}} = \frac{1}{\sqrt{2n}} .$$

(f)— *2p* – Finally, the 95% confidence interval for ψ is

$$\widehat{\psi}_n \pm 1.96 \widehat{\mathsf{se}}_n(\widehat{\Psi}_n) = \log(\widehat{\sigma}_n) \pm 1.96 \frac{1}{\sqrt{2n}}.$$

(g)— *1p* – Since $n = 110$ and $\widehat{\sigma}_n = 12.4$, the $(1-\alpha) = 95\%$ confidence interval for ψ is

$$\log(\widehat{\sigma}_n) \pm 1.96 \frac{1}{\sqrt{2n}} = \log(12.4) \pm 1.96 \frac{1}{\sqrt{2\times 110}} = 0.8754 \pm 0.132 = [0.7433, 1.0076]$$

Since $(1-\alpha) = 95\%$ confidence interval for ψ does not contain 10.0, we reject $H_0 : \psi = 10.0$ in favour of $H_1 : \psi \neq 10.0$ using a size $\alpha = 0.05$ Wald test.

(h)— *1p* – The 95% confidence interval for σ based on $n = 110$ samples and a sample standard deviation of 12.4 is:

$$\widehat{\sigma}_n \pm 1.96 \frac{\widehat{\sigma}_n}{\sqrt{2n}} = 12.4 \pm 1.96 \frac{12.4}{\sqrt{2\times 110}} = [10.76, 14.04]$$

(i)— *1p* – Since the 95% confidence interval for σ does contain 10.0, we fail to reject the null hypothesis $H_0 : \sigma = 10.0$ versus $H_1 : \sigma \neq 10.0$ using a size $\alpha = 0.05$ Wald test.

Answer (Ex. 5) — — *1p* – First set up what you need to show:
We need to show that: $P_\lambda(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n|\bar{X}_n)$ is independent of λ for any $(x_1, x_2, \ldots, x_n) \in \{0, 1, 2, \ldots\}^n$ and any $\bar{x}_n \geq 0$.

— *1p* – Realising the following equality
$P_\lambda(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n|\bar{X}_n) = \frac{P_\lambda(X_1=x_1,X_2=x_2,\ldots,X_n=x_n,\bar{X}_n=\bar{x}_n)}{P_\lambda(\bar{X}_n=\bar{x}_n)}$

— *1p* – Realising $n\bar{X}_n = \sum_{i=1}^n X_i \sim Poisson(n\lambda)$ and noting that
$P_\lambda(n\bar{X}_n = n\bar{x}_n) = e^{-n\lambda}\frac{(n\lambda)^{\sum_{i=1}^n x_i}}{(\sum_{i=1}^n x_i)!} = e^{-n\lambda}\frac{(n\lambda)^k}{k!}, \quad k := n\bar{x}_n$

— *1p* – Realising the last equality below, with say $k = n\bar{x}_n \geq 0$

$$\frac{P_\lambda(X_1=x_1,X_2=x_2,\ldots,X_n=x_n,\bar{X}_n=\bar{x}_n)}{P_\lambda(\bar{X}_n=\bar{x}_n)} = \frac{P_\lambda(X_1=x_1,X_2=x_2,\ldots,X_n=x_n,n\bar{X}_n=k)}{P_\lambda(n\bar{X}_n=k)} = \frac{P_\lambda(X_1=x_1,X_2=x_2,\ldots,X_n=x_n)}{P_\lambda(n\bar{X}_n=k)}$$

— *1p* – Realising the following equality

$$\frac{P_\lambda(X_1=x_1,X_2=x_2,\ldots,X_n=x_n)}{P_\lambda(n\bar{X}_n=k)} = \frac{\prod_{i=1}^n \left(e^{-\lambda}\frac{\lambda^{x_i}}{x_i!}\right)}{e^{-n\lambda}\frac{(n\lambda)^k}{k!}} = \frac{e^{-n\lambda}\frac{\lambda^k}{\prod_{i=1}^n x_i!}}{e^{-n\lambda}\frac{(n\lambda)^k}{k!}}$$

— *1p* – Finally, showing that the conditional probability is independent of parameter explicitly:

$$P_\lambda(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n|\bar{X}_n) = \frac{k!}{n^k \prod_{i=1}^n x_i!} = (\textstyle\sum_{i=1}^n x_i)!\Big(n^{(\sum_{i=1}^n x_i)}\prod_{i=1}^n x_i!\Big)^{-1}$$

Answer (Ex. 6) — (a)– *1p* — For writing what needs to be shown:
Fix any x. Bias is the expected value of the estimator, so we need to show that $E(\widehat{F}_n(x)) - F(x) \to 0$ as $n \to \infty$.

And *1p* — For showing:
$E(\widehat{F}_n(x)) = E(n^{-1}\sum_{i=1}^n \mathbf{1}(X_i \leq x)) = n^{-1}\sum_{i=1}^n E(\mathbf{1}(X_i \leq x))$
$= n^{-1}\sum_{i=1}^n P(X_i \leq x) = n^{-1}\sum_{i=1}^n F(x) = F(x)$.

Thus $\widehat{F}_n(x)$ is an unbiased estimator of $F(x)$ for any x.

(b)– *2p* — Realise that $Y_i := \mathbf{1}(X_i \leq x) \sim Bernoulli(\theta = F(x))$ RV and use IID sum of Bernoulli RVs

$$V(\widehat{F}_n(x)) = V\left(n^{-1}\sum_{i=1}^{n} Y_i\right) = V\left(\sum_{i=1}^{n} n^{-1} Y_i\right) = n^{-2}\sum_{i=1}^{n} V(X)$$

Since variance of a $Bernoulli(\theta)$ RV is $\theta(1-\theta)$, with $\theta = F(x)$ for us now, we get the following continuation of equalities from the previous line:

$$n^{-2}\sum_{i=1}^{n} V(X) = n^{-2} n V(X) = n^{-1} V(X) = \frac{V(X)}{n} = \frac{F(x)(1-F(x))}{n}$$

(c)– *1p* — Realizing IID likelihood is constant over the permutations of the m data points:

$$X_1, \ldots, X_m \overset{IID}{\sim} F \implies \text{,i.e., implies, the following equality:}$$

$$P(X_1 = x_1, \ldots, X_m = x_m) = \prod_{i=1}^{m} P(X_i = x_i) = \prod_{i=1}^{m} P(X_i = x_{\pi(i)})$$

And since there are $m!$ possible permutations and each of them is equally likely gives $P(X_1 = x_1, X_2 = x_2, \ldots, X_m = x_m) = P(X_1 = x_{\pi(1)}, X_2 = x_{\pi(2)}, \ldots, X_m = x_{\pi(m)}) = 1/m!$. [To properly understand what is going on here, you may want to recall the scribed example from last couple weeks' lectures on (nonparametric and exact) permutation test with 2 samples from one and 1 sample from a possibly different population with the probabilities being equally 1/6 for all $6 = (2+1)!$ permutations.]

Feedback from "raazozone": (1) Less time will be spent on Python soon as all maths students will have a full semester of Python as requirement in a year or so [thanks for the student who gace feedback formally!], (2) Raaz, i.e. I, will most likely be teaching Probability 1 and Inference Theory 1 next year and therefore the contents will be tightly integrated, (3) the current vision is to meet my legal obligations and be able to teach the course in Swedish by 2021, (4) For those who are too busy to attend scheduled course meetings — note that $L :=$ "Lectures", $C :=$ "Computer-laboratories", $P :=$ "Problem-sessions" when Raazesh Sainudiin teaches Inference Theory 1 will remain an amalgamation of face-to-face communications/learning/teaching at the scheduled course meetings using the mediums of black-board and portable-computers (available in UU's computer labs for those without such a device) and NOT be broken into pre-made partitions in time for $\{L, C, P\}$ for the commuting conveniences of some — one of your colleagues had kindly allowed me to take some images and post them online here: `https://github.com/lamastex/scalable-data-science/raw/master/_infty/2018/01/scribed/arch/soFar.pdf` (I am aware that the image quality is not the best, but that's the brakes from me! I am happy to remove them from the course URL completely to be more precise). Mathematics is not a spectator sport, it's an art-form or an unnatural language that is learned by doing and communicating in face-to-face transmissions, much like you would learn the basics of sculpting, pottery, of other fine-arts from someone who has more experience in the area.