Statistics – HW 2 – Solution

1) Direct computation.

$$F_\theta(t) = 1 - e^{-\theta t}$$

2) Solve for Q_1, Q_2, Q_3 such that

$F_\theta(Q_1) = 0.25 \qquad 1 - e^{-\theta Q_1} = 0.25$

$F_\theta(Q_2) = 0.5 \qquad 1 - e^{-\theta Q_2} = 0.5$

$F_\theta(Q_3) = 0.75 \qquad 1 - e^{-\theta Q_3} = 0.75$

$$\boxed{\begin{aligned} Q_1 &= \frac{-\ln(0.75)}{\theta} \\ Q_2 &= \frac{-\ln(0.5)}{\theta} \\ Q_3 &= \frac{-\ln(0.25)}{\theta} \end{aligned}}$$

3) We get

$$Q_3 - Q_1 = \frac{1}{\theta}\left(-\ln(0.25) + \ln(0.75)\right)$$

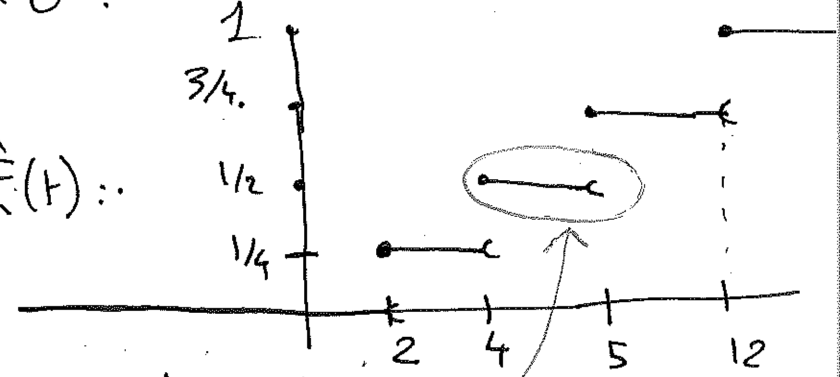
$\ln$ is increasing so $\ln(0.75) - \ln(0.25) \geq 0$

$Q_3 - Q_1$ is decreasing with θ.

4) Take $n = 4$

$X_1 = 2 \quad X_2 = 4 \quad X_3 = 5 \quad X_4 = 12$

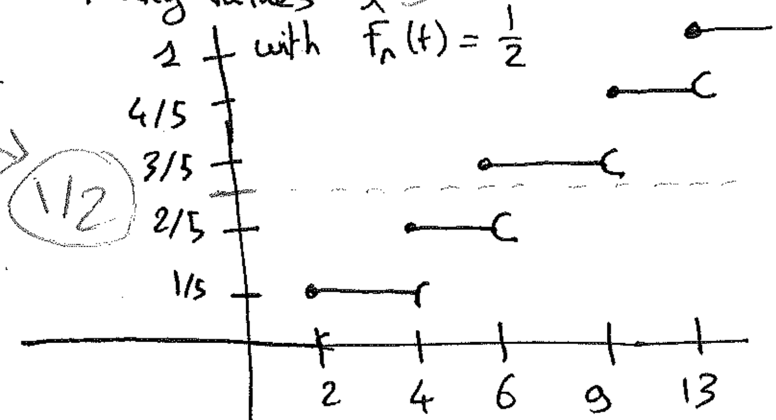
$\hat{F}_n(t)$:



Many values with $\hat{F}_n(t) = \frac{1}{2}$

$n = 5$

No value equal to $1/2$



$X_1 = 2 \quad X_2 = 4 \quad X_3 = 6 \quad X_4 = 9 \quad X_5 = 13$

In general, if $n = 2k$ (is even),

then $\hat{F}_n(t) = \frac{1}{2}$ between the k^{th} and $k+1^{th}$ value of the data set.

So, many points.

if $n = 2k+1$ (is odd)

$\hat{F}_n(t)$ jumps from $\frac{k}{2k+1} < \frac{1}{2}$) no value $= 1/2$.

to $\frac{k+1}{2k+1} > \frac{1}{2}$

5] • For n even $(n = 2k)$, we could take

* the k^{th} value ; * the $k+1^{th}$ value

* the midpoint between these two values?

• For n odd $(n = 2k+1)$ we could take

* the k^{th} value, the $k+1^{th}$ value

* the midpoint between these two values?

My preferred choice. I believe this reduces the bias, i.e.

$\mathbb{E}[\hat{M}_n] - M$ is smaller than for the other choices.

6] If $M + \varepsilon \leq \hat{M}_n$, since $\hat{F}_n$ is non-decreasing, we have

$\hat{F}_n(M+\varepsilon) \leq \hat{F}_n(\hat{M}_n)$.

On the other hand, we guaranteed that $|\hat{F}_n(\hat{M}_n) - \frac{1}{2}| \leq \frac{1}{n}$, so in particular

$\hat{F}_n(\hat{M}_n) \leq \frac{1}{2} + \frac{1}{n}$, hence

if $M + \varepsilon \leq \hat{M}_n$, then $\hat{F}_n(M+\varepsilon) \leq \frac{1}{2} + \frac{1}{n}$. So

$\mathbb{P}(M + \varepsilon \leq \hat{M}_n) \leq \mathbb{P}(\hat{F}_n(M+\varepsilon) \leq \frac{1}{2} + \frac{1}{n})$.

7] We know that for any t, $\hat{F}_n(t)$ is a consistent estimator of $F(t)$, which means $\hat{F}_n(t) \xrightarrow[n\to\infty]{\mathbb{P}} F(t)$ (in probability).

8] We want to prove that $\hat{M}_n \xrightarrow[n\to\infty]{\mathbb{P}} M$, ie.

$$\forall \varepsilon > 0, \quad \lim_{n\to\infty} \mathbb{P}\left(|\hat{M}_n - M| \geq \varepsilon \right) = 0 .$$

a) Let us write $\mathbb{P}\left(|\hat{M}_n - M| \geq \varepsilon \right) \leq \mathbb{P}\left(\hat{M}_n - M \geq \varepsilon \right) + \mathbb{P}(\hat{M}_n - M \leq -\varepsilon)$ and focus on the term $\mathbb{P}(\hat{M}_n - M \geq \varepsilon)$, the other one is treated similarly.

We want to show

$$\forall \varepsilon > 0, \quad \lim_{n\to\infty} \mathbb{P}\left(\hat{M}_n - M \geq \varepsilon \right) = 0 .$$

b) By Q.6], we know $\mathbb{P}\left(\hat{M}_n - M \geq \varepsilon \right) \leq \mathbb{P}\left(\hat{F}_n(M+\varepsilon) \leq \frac{1}{2} + \frac{1}{n} \right)$, so it is enough to prove that

$$\forall \varepsilon > 0, \quad \lim_{n\to\infty} \mathbb{P}\left(\hat{F}_n(M+\varepsilon) \leq \frac{1}{2} + \frac{1}{n} \right) = 0 .$$ Fix $\varepsilon > 0$.

c) By Q.7], we know that $\hat{F}_n(M+\varepsilon) \xrightarrow[n\to\infty]{\mathbb{P}} F(M+\varepsilon)$, and by assumption $F(M+\varepsilon)$ is strictly larger than $\frac{1}{2}$, because F is assumed to be strictly increasing. Let us write $F(M+\varepsilon) = \frac{1}{2} + \gamma$, with $\gamma > 0$. The convergence $\hat{F}_n(M+\varepsilon) \xrightarrow[n\to\infty]{\mathbb{P}} F(M+\varepsilon)$ can be written

$$\forall \delta > 0, \quad \lim_{n\to\infty} \mathbb{P}\left(\left|\hat{F}_n(M+\varepsilon) - F(M+\varepsilon)\right| \geq \delta \right) = 0 .$$

In particular, taking $\delta = \frac{\gamma}{2}$, we have

$$\lim_{n\to\infty} \mathbb{P}\left(\hat{F}_n(M+\varepsilon) \leq F(M+\varepsilon) - \delta \right) = 0$$

so $$\lim_{n\to\infty} \mathbb{P}\left(\hat{F}_n(M+\varepsilon) \leq \frac{1}{2} + \gamma - \frac{\gamma}{2} \right) = 0$$

so $$\lim_{n\to\infty} \mathbb{P}\left(\hat{F}_n(M+\varepsilon) \leq \frac{1}{2} + \frac{\gamma}{2} \right) = 0 .$$

For n large enough, $\frac{1}{n} \leq \frac{\gamma}{2}$, so $\lim_{n\to\infty} \mathbb{P}\left(\hat{F}_n(M+\varepsilon) \leq \frac{1}{2} + \frac{1}{n} \right) \leq \lim_{n\to\infty} \mathbb{P}\left(\hat{F}_n(M+\varepsilon) \leq \frac{1}{2} + \frac{\gamma}{2} \right)$

The right-hand side is $=0$, so we have proven, as desired,

$$\lim_{n\to\infty} \mathbb{P}\left(\hat{F}_n(\tilde{M}_n+\varepsilon) \leq \frac{1}{2}+\frac{1}{n} \right) = 0,$$

which concludes the proof.