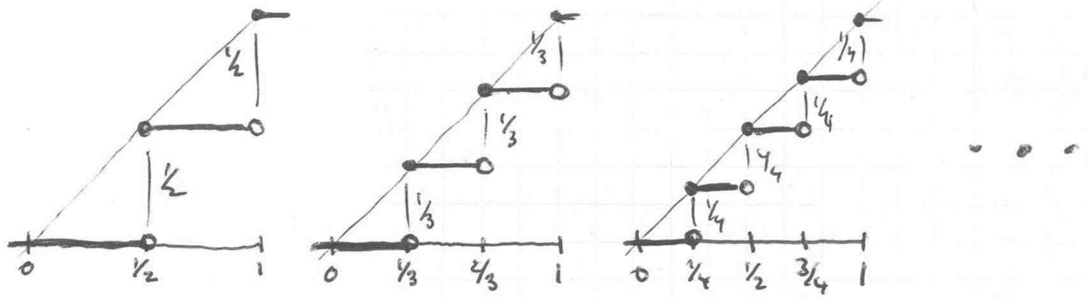
$\boxed{1.1}$ Let X_n be distributed s.t. $P(X_n = \frac{i}{n}) = \frac{1}{n}$ for all $1 \le i \le n$. Show that $X_n \rightsquigarrow X \sim U[0,1]$.

prf: The distribution functions for X_n are:

$$P(X_n \le x) = \sum_{\{i \,:\, i/n \le x\}} P\left(X_n = \frac{i}{n}\right)$$

$$= \frac{1}{n} \lfloor nx \rfloor \quad (\text{"entier" of } nx)$$



Either by the above graphs or by calculus for entiers, we see that for every $x \in \mathbb{R}$,

$$P(X \le x) - \frac{1}{n} \le P(X_n \le x) \le P(X \le x)$$

Where $P(X \le x) = x$ if $x \in [0,1]$, $P(X \le x) = 0$ for all $x < 0$ and $P(X \le x) = 1$ for all $x > 1$.
We conclude that $X_n \rightsquigarrow U[0,1]$.

$\boxed{1.25}$ Suppose that $\sqrt{n}(T_n - \theta) \rightsquigarrow Z$.

Let $\epsilon > 0$ be given, consider

$$P(|T_n - \theta| > \epsilon) = 1 - P(-\epsilon \le T_n - \theta \le \epsilon)$$

$$= 1 - P(-\sqrt{n}\,\epsilon \le \sqrt{n}(T_n - \theta) \le \sqrt{n}\,\epsilon)$$

Assuming for simplicity that Z_n has a continuous distribution:

$$= 1 - P(-\sqrt{n}\,\epsilon \le \sqrt{n}(T_n - \theta) \le \sqrt{n}\,\epsilon)$$

$$= P(\sqrt{n}(T_n - \theta) \ge \sqrt{n}\,\epsilon)$$
$$\quad + P(\sqrt{n}(T_n - \theta) \le -\sqrt{n}\,\epsilon)$$

For any $M > 0$, there exists an N s.t. for all $n \ge N$, $\sqrt{n}\,\epsilon > M$, so that

$$\le P(\sqrt{n}(T_n - \theta) \ge M) + P(\sqrt{n}(T_n - \theta) \le -M)$$

$$\longrightarrow P(Z \ge M) + P(Z \le -M)$$

$$= \cancel{P(|Z| \ge M)} \; P(|Z| > M)$$

For every $\delta > 0$, there is an $M > 0$ s.t. $P(|Z| > M) < \delta$.

$\boxed{1.11}$ Find an example of sequences (X_n) and (Y_n) that converge weakly: $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$, but (X_n, Y_n) does not converge.

<u>answer</u>: we shall have to make a dependent choice for (X_n, Y_n), because if they are independent, then convergence $(X_n, Y_n) \xrightarrow{P} (X, Y)$.

Consider $X_n \sim N(0,1)$ (for all $n \geq 1$) and $Y_n = (-1)^n X_n$

Then $X_n \xrightarrow{P} X \sim N(0,1)$ and since $X \stackrel{\mathcal{D}}{=} -X$, $Y_n \xrightarrow{P} X \sim N(0,1)$ as well.

Let $x, y \in \mathbb{R}$ be given; consider

$$P(X_n \leq x, Y_n \leq y)$$

$$= \begin{cases} P(X \leq x, X \leq y) & \text{for } n \in 2\mathbb{N} \\ P(X \leq x, -X \leq y) & \text{for } n \in 2\mathbb{N} - 1 \end{cases}$$

$$= \begin{cases} P(X \le \min(x,y)), & \text{for } n \in 2\mathbb{N} \\ P(X \le x,\ X \ge -y), & \text{for } n \in 2\mathbb{N}-1 \end{cases}$$

$$= \begin{cases} P(X \le \min(x,y)), & \text{for } n \in 2\mathbb{N} \\ P(X \in [-y, x]), & \text{for } n \in 2\mathbb{N}-1 \end{cases}$$

So if we take $x < -y$, $[-y, x] = \emptyset$ and the bottom line is 0. However, for any $(x,y) \in \mathbb{R}^2$, $P(X \le \min(x,y)) = \Phi(\min(x,y)) > 0$. Hence

$$\limsup_{n\to\infty} P(X_n \le x, Y_n \le y) > 0$$
$$\liminf_{n\to\infty} P(X_n \le x, Y_n < y) = 0 \qquad \square$$

$\boxed{1.20}$ Suppose $X_n \sim N(\mu_n, \sigma_n^2)$.

(i) If $\sigma_n = 1$ for all $n \geq 1$, show that X_n is unif. tight $\iff$ $\mu_n = O(1)$.

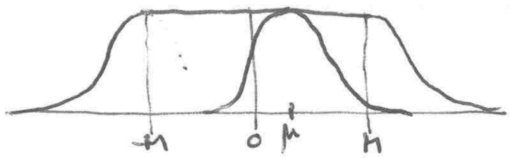
'$\Leftarrow$' There exists a constant $M > 0$ s.t.

$$-M \leq \inf \mu_n \leq \sup \mu_n \leq M,$$

Note that for any $\mu \in [-M, M]$ & $X \sim N(\mu, 1)$

$$P_\mu(X \geq x) \leq P_M(X \geq x)$$

$$P_\mu(X \leq -x) \leq P_{-M}(X \leq -x)$$



for any $x \geq M$. Therefore:

$$\sup_{n \geq 1} P_{\mu_n}(X \geq x) = \sup_{n \geq 1} P(X_n \geq x) \leq P_M(X \geq x)$$

$$\sup_{n \geq 1} P_{\mu_n}(X \leq -x) = \sup_{n \geq 1} P(X_n \leq -x) \leq P_{-M}(X \leq -x)$$

Because $N(M, 1)$ and $N(-M, 1)$ are tight for every $\epsilon>0$ there exists an $L>0$ s.t.

$$P_M(X \geq L) \vee P_{-M}(X \leq -L) < \epsilon$$

which implies that,

$$\sup_{n\geq 1} P(|X_n| \geq L) < \epsilon.$$

'$\Rightarrow$' Let $0<\epsilon<\frac{1}{2}$ be given. There exists an $L>0$ s.t. $P(|X_n| \geq L) < \epsilon$ for all $n \geq 1$. It follows that $|\mu_n| \leq L$ for all $n \geq 1$. So $\mu_n = O(1)$.

(ii) Let $\mu_n = 0$ for all $n \geq 1$. For which (σ_n) is (X_n) unif. tight? For all $L>0$,

$$P(|X_n| \geq L) = 2\int_{-\infty}^{-L} \frac{1}{\sqrt{2\pi}\,\sigma_n} \exp\left(\frac{-x^2}{2\sigma_n^2}\right) dx$$

$$= 2\int_{-\infty}^{-L/\sigma_n} \frac{1}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}y^2\right) dy$$

$$= 2\Phi\left(-\frac{L}{\sigma_n}\right)$$

Let $M>0$ be given. and suppose ...

Suppose that $|\sigma_n| \le M$ f.l.e.n. Then,

$$\sup_{n \ge N} P(|X_n| \ge L) \le 2\,\Phi\left(-\frac{L}{M}\right) \quad (<\epsilon)$$

so for any $\epsilon > 0$, we can find $L > 0$ s.t.

Suppose that for any $M > 0$ there is an $n \ge 1$ s.t. $\sigma_n \ge M$. Then, for that n,

$$P(|X_n| \ge L) \ge 2\,P\left(-\frac{L}{M}\right)$$

Indeed, $\limsup \sigma_n = \infty$, so

$$\liminf_{n \to \infty} P(|X_n| \ge L) \ge 2 \liminf_{n \to \infty} \Phi\left(-\frac{L}{\sigma_n}\right)$$

$$= 2 \cdot \frac{1}{2} = 1.$$

So we can conclude that

$$X_n \text{ unif. tight} \iff \sigma_n = O(1).$$

$\square$

2.3 Determine the covariance matrix for a random vector $X \sim$ Multinomial $(n; p_1, \dots, p_k)$.

Consider first the case that $n=1$, i.e. we have $Y = (Y_1, \dots, Y_k)$ with only one $Y_i = 1$ and the rest $Y_j = 0$ $(j \neq i)$, with

$$P(Y=y) = p_1^{y_1} p_2^{y_2} \cdots p_k^{y_k}$$

$(\sum p_i = 1,\ p_i \geq 0)$. Therefore, $EY = p$. The covariance matrix follows from the observation that $=P(Y_i=1)$, $Y_i Y_j = \delta_{ij}$, ~~and~~ so that $E(Y_i Y_j = 1 | Y_i = 1) = \delta_{ij}$. Therefore, $EY_iY_j = E(Y_iY_j | Y_i = 1)\, P(Y_i=1) = \delta_{ij} p_i$

and

$$\operatorname{cov}(Y_i, Y_j) = E(Y_i - EY_i)(Y_j - EY_j)$$

$$= EY_iY_j - 2EY_i\, EY_j + EY_i\, EY_j$$

$$= \delta_{ij} p_i - p_i p_j$$

For $X \sim$ Multinomial $(n; p_1, \dots, p_k)$, observe that X is the sum of n independent random vectors $Y^{(1)}, Y^{(2)}, \dots, Y^{(n)}$ each of which is multinomial $(1; p_1, \dots, p_k)$.

And if $X \perp Y$, $\operatorname{Cov}(X+Y) = \operatorname{Cov}(X) + \operatorname{Cov}(Y)$, so

$$\operatorname{Cov}(X_i, X_j) = \operatorname{Cov}(Y_i^{(1)} + \dots + Y_i^{(n)}, Y_j^{(1)} + \dots + Y_j^{(n)})$$

$$= \operatorname{Cov}(Y_i^{(1)}, Y_j^{(1)}) + \dots + \operatorname{Cov}(Y_i^{(n)}, Y_j^{(n)})$$

$$= n(\delta_{ij} p_i - p_i p_j)$$

$\boxed{2.17}$ $X \sim N_k(0, \mathbb{1}_k)$, $H_0 \subset H \subset \mathbb{R}^k$ lin subsp.

(i) Show that $\Pi_0 X$ and $\Pi^\perp X = (1-\Pi)X$ are independent. (Π_0, Π orth. proj. H_0, H)

Prf Consider the random vector $Y \in \mathbb{R}^{2k}$

$$Y = (\Pi_0 X, (1-\Pi)X)$$

which has a multivariate normal distr.
The mixed covariances are

$$\mathrm{Cov}(\Pi_0 X, (1-\Pi)X) = 0$$

because $H_0 \perp H^\perp$. Combining, we see that $\Pi_0 X$ and $(1-\Pi)X$ are independent
So the multivariate normal distribution for Y is a product $N_k(0, \Sigma_0) \times N_k(0, \Sigma')$ where Σ_0 and Σ' may be singular depending on the dimensions of H_0 and H.

(ii) The joint distribution of $(\|\Pi_0 X\|^2, \|(1-\Pi)X\|^2)$ is $\chi^2_{\ell_0} \times \chi^2_{\ell'}$ where ℓ_0 is the dim. of H_0 and $\ell' = k - \dim(H)$.

$\boxed{3.i}$ Let $\hat{\lambda}_n$ be the MLE for the parameter λ of the $\mathrm{Exp}(\lambda)$-distributed iid $X_1, \dots, X_n$.

<u>NB</u> Use that $\mathrm{Exp}(\lambda)$ has density

$$p_\lambda(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \mathbb{1}\{x \geq 0\}$$

and satisfies $E_\lambda X = \lambda$, $\mathrm{Var}_\lambda(X) = \lambda^2$.

First we need the MLE for λ:

$$L_n(\bar{X}_n; \lambda) = \prod_{i=1}^n p_\lambda(X_i) = \lambda^{-n} e^{-\lambda^{-1} \sum_{i=1}^n X_i} \mathbb{1}\{X_{(1)} \geq 0\}$$

$$= \lambda^{-n} e^{-n\lambda^{-1} \bar{X}_n} \quad (a.s.)\left(P_\lambda(X_i \geq 0) = 1 \right)$$

Therefore, by taking the log and $\partial/\partial\lambda$:

$$\frac{\partial}{\partial\lambda} \log L_n = \frac{\partial}{\partial\lambda}\left(-n \log \lambda - \frac{n}{\lambda} \bar{X}_n \right)$$

$$= -\frac{n}{\lambda} + \frac{n}{\lambda^2} \bar{X}_n \quad \Rightarrow \quad \hat{\lambda}_n(\underline{X}_n) = \bar{X}_n.$$

(i) Therefore,

$$\sqrt{n}(\hat{\lambda}_n - \lambda) = \sqrt{n}(\bar{X}_n - E_\lambda X) \xrightarrow{\lambda} N(0, \mathrm{Var}_\lambda(X))$$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \overset{\|}{N(0, \lambda^2)}$$

(iii) Consider the following def. of asymp. CI.

$$P_\lambda^n(\lambda \in \hat{I}_{n,\alpha}) \longrightarrow 1-\alpha.$$

So the statistic here is $\hat{I}_{n,\alpha} = [\hat{\ell}_{n,\alpha}, \hat{u}_{n,\alpha}]$ and the usual coverage prob. $1-\alpha$ for the (level-α) CI is now a limit.

In the present case, we know that

$$P_\lambda^n\left(x \leq \frac{\sqrt{n}(\hat{\lambda}_n - \lambda)}{\lambda} \leq y \right)$$

$$= P_\lambda^n\left(\sqrt{n}\left(\frac{\hat{\lambda}_n}{\lambda} - 1\right) \leq y \right) - P_\lambda^n(\ldots \leq x)$$

$$\longrightarrow \Phi(y) - \Phi(x).$$

So, $$P_\lambda^n\left(x < \sqrt{n}\left(\frac{\hat{\lambda}_n}{\lambda} - 1\right) \leq y \right)$$

$$= P_\lambda^n\left(\frac{x}{\sqrt{n}} < \frac{\hat{\lambda}_n}{\lambda} - 1 \leq \frac{y}{\sqrt{n}} \right)$$

$$= P_\lambda^n\left(\frac{x}{\sqrt{n}} + 1 < \frac{\hat{\lambda}_n}{\lambda} \leq \frac{y}{\sqrt{n}} + 1 \right)$$

$$= P_\lambda^n\left(\hat{\lambda}_n\left(\frac{y}{\sqrt{n}} + 1\right)^{-1} \leq \lambda < \hat{\lambda}_n\left(\frac{x}{\sqrt{n}} + 1\right) \right)$$

Now take $x = -\xi_{\alpha/2}$, $y = \xi_{\alpha/2}$: $\longrightarrow 1-\alpha$

So $$\hat{I}_{n,\alpha} = \left[\hat{\lambda}_n\left(\frac{\xi_{\alpha/2}}{\sqrt{n}} + 1\right)^{-1}, \hat{\lambda}_n\left(1 - \frac{\xi_{\alpha/2}}{\sqrt{n}}\right)^{-1} \right]$$

$\boxed{3.11}$ let $X_1, \ldots, X_n$ be iid P with $EX = \mu$, $\mathrm{Var} X = 1$

Q: Find constants a_n, b_n s.t. $a_n(\bar{X}_n^2 - b_n)$ converges weakly, in cases $\mu = 0$ and $\mu \neq 0$.

Since $\mathrm{Var}(X) < \infty$, the CLT says that

$$\sqrt{n}(\bar{X}_n - \mu) \overset{P}{\rightsquigarrow} N(0, \mathrm{Var}(X)) = N(0,1).$$

Now take $\phi(x) = x^2$ on $\mathbb{R}$. Then $\phi'(x) = 2x$

and $\sqrt{n}(\phi(\bar{X}_n) - \phi(\mu)) \overset{P}{\rightsquigarrow} N(0, \phi'(\mu)^2)$.

In case that $\mu \neq 0$, $\phi'(\mu)^2 = 4\mu^2 > 0$. So then,

$$\sqrt{n}(\bar{X}_n^2 - \mu^2) \overset{P}{\rightsquigarrow} N(0, 4\mu^2). \quad (a_n = \sqrt{n}, b_n = \mu)$$

If $\mu = 0$ we take a different approach:

$$\sqrt{n}\bar{X}_n \overset{P}{\rightsquigarrow} N(0,1)$$

so by cont. mapping, $\phi(x) = x^2$,

$$n\bar{X}_n^2 \overset{P}{\rightsquigarrow} \chi_1^2 \qquad (a_n = n, b_n = 0)$$

because if $Z \sim N(0,1)$, then $\phi(Z) = Z^2 \sim \chi_1^2$. $\square$

$\boxed{3.16}$ Let $X_1, \dots, X_n \sim$ iid-P_θ where P_θ is from a 1-dim exp. family with densities

$$p_\theta(x) = c(\theta)\, h(x)\, e^{\theta t(x)}$$

(with known functions h, t).

The natural parameter set for this model,

$$\Theta = \left\{ \theta \in \mathbb{R} : \int h(x) e^{\theta t(x)} dx < \infty \right\}$$

(i) Show that Θ is an interval.

Note that $\theta \mapsto e^{\theta t(x)}$ is convex ($\forall_x$). So for every $\theta_1 < \theta_2$ in Θ and $\lambda \in [0,1]$,

$$\int h(x) e^{(\lambda\theta_1 + (1-\lambda)\theta_2)\, t(x)} dx$$

$$\leq \int h(x) \left(\lambda e^{\theta_1 t(x)} + (1-\lambda) e^{\theta_2 t(x)} \right) dx$$

$$= \lambda \int h(x) e^{\theta_1 t(x)} dx + (1-\lambda) \int h(x) e^{\theta_2 t(x)} dx$$

So if $\int h(x) e^{\theta t(x)} dx < \infty$ in θ_1 and θ_2, then $\int h(x) e^{\theta t(x)} dx < \infty$ for all $\theta \in [\theta_1, \theta_2]$.

Hence, Θ is interval, (non-empty if $<\infty$ for some θ)

[This is a manifestation of something far more general: if $f: X \to Y$ is convex, then $\{x \in X : f(x) < y\}$ are all convex in X.]

(ii) Show that the MLE $\hat{\theta}_n$ is also a moment estimator.

$$L_n(\theta, \underline{X}_n) = \prod_{i=1}^{n} c(\theta)\, h(X_i)\, e^{\theta t(X_i)}$$

$$= c(\theta)^n\, e^{\theta \sum_{i=1}^{n} t(X_i)} \left(\prod_{i=1}^{n} h(X_i) \right)$$

where $c(\theta) = \left(\int h(x) e^{\theta t(x)} dx \right)^{-1}$. Take a log and a derivative:

$$\dot{l}_\theta(x) = \frac{d}{d\theta} \log p_\theta(x) = \frac{d}{d\theta} \left(\log c(\theta) + \theta t(x) + .. \right)$$

$$= c'(\theta)\, c^{-1}(\theta) + t(X)$$

Assuming that $\frac{d}{d\theta} \int h(x)\, e^{\theta t(x)} dx \overset{(A)}{=} \int \frac{\partial}{\partial \theta} \left(h e^{\theta t} \right) dx$

we find that,

$$\frac{d}{d\theta} \log c(\theta) = - \frac{d}{d\theta} \log \int h e^{\theta t} dx = -c(\theta) \int h(x) t(x) e^{\theta t} dx$$

so the score equation becomes:

$$\sum_i t(X_i) = n\, c(\theta) \int h(x)\, t(x)\, e^{\theta t(x)} dx$$

$$\mathbb{P}_n t(X) = \mathbb{E}_\theta t(X). \quad \text{(moment equation)}$$

The solution to the MLE problem is moment est.

$\boxed{3.19}$ Suppose $\phi: \mathbb{R} \to \mathbb{R}$ is 2x cont. diff. at θ with $\phi'(\theta) = 0$, $\phi''(\theta) \neq 0$. Let $\sqrt{n}(T_n - \theta) \overset{\theta}{\rightsquigarrow} N(0,1)$

(i) Show that $\sqrt{n}\left(\phi(T_n) - \phi(\theta)\right) \xrightarrow{P_\theta} 0$

(ii) Show that $n\left(\phi(T_n) - \phi(\theta)\right) \overset{P_\theta}{\rightsquigarrow} \chi^2_1$

We shall prove both (i) and (ii) simultaneously

prf Because $\sqrt{n}(T_n - \theta) \overset{P_\theta}{\rightsquigarrow} N(0,1)$, we have $n(T_n - \theta)^2 \overset{P_\theta}{\rightsquigarrow} \chi^2_1$ by cont. mapping.

Note that this implies also that $T_n \xrightarrow{P_\theta} \theta$.

By assumption, ϕ has a Taylor expansion

$$\phi(\theta + h) = \phi(\theta) + \tfrac{1}{2} h^2 \phi''(\theta) + o(h^2),$$

(as $h \to 0$). So the function,

$$g(h) = \begin{cases} \dfrac{\phi(\theta+h) - \phi(\theta) - \frac{1}{2}h^2 \phi''(\theta)}{h^2}, & \text{if } h \neq 0 \\ 0, & \text{if } h = 0 \end{cases}$$

is continuous at $h=0$. Hence, using $T_n - \theta \xrightarrow{P_\theta} 0$ and the continuous mapping theorem,

$$g(T_n - \theta) \xrightarrow{P_\theta} 0.$$

Using that $n(T_n-\theta)^2 \overset{P_\theta}{\rightsquigarrow} \chi^2_1$, we get

$$n(T_n-\theta)^2 g(T_n-\theta) \xrightarrow{P_\theta} 0$$

From Slutsky's lemma, so we see,

$$\left| n(\phi(T_n)-\phi(\theta)) - \frac{n}{2}\phi''(\theta)(T_n-\theta)^2 \right| \xrightarrow{P_\theta} 0.$$

The second term has $\frac{1}{2}\phi''(\theta)\chi^2_1$ as it's weak limit. Consequently, (by Slutsky)

$$n(\phi(T_n)-\phi(\theta)) \overset{P_\theta}{\rightsquigarrow} \frac{1}{2}\phi''(\theta) Z^2$$

Where $Z \sim N(0,1)$, so that $Z^2 \sim \chi^2_1$.

Therefore, again using Slutsky,

$$\sqrt{n}(\phi(T_n)-\phi(\theta)) = \frac{1}{\sqrt{n}} n(\phi(T_n)-\phi(\theta))$$

converges $\overset{\text{weakly}}{\text{to}}$ $0 \cdot \chi^2_1 = 0$, so also

$$\sqrt{n}(\phi(T_n)-\phi(\theta)) \xrightarrow{P_\theta} 0.$$

$\boxed{4.2}$ Let $X_1, \dots, X_n \sim$ iid $-P$, where P has a density $p(x) > 0$ $(\forall_{x \in \mathbb{R}})$ with some $\theta_0 \in \mathbb{R}$ s.t. $p(\theta_0 + a) = p_0(\theta_0 - a)$ for all $a \geq 0$.

Show that the Huber estimator for location is consistent for θ_0.

The Huber estimator $\hat{\theta}_n$ solves $\mathbb{P}_n \psi(X-\theta) = 0$ with

$$\psi(y) = \begin{cases} -k & \text{if } y < -k \\ y & \text{if } |y| \leq k \\ k & \text{if } y > k \end{cases}$$

Define $\Psi(\theta) = P\psi(X-\theta)$. Note that $\Psi_n(\theta)$ satisfies

$$\Psi_n(\theta) = \mathbb{P}_n \psi(X-\theta) \xrightarrow{as.} \Psi(\theta)$$

under any P, because $|\psi| \leq k$ (so LLN).

Because ψ is increasing $\theta \mapsto \psi(x-\theta)$ is decreasing, so any sum $\sum_{i=1}^n \psi(x_i - \theta)$ is decreasing. Hence $\theta \mapsto \Psi_n(\theta)$ is non-increasing. By the same argument, neither is $\Psi(\theta)$.

Consider $\Psi(\theta_0 + a)$ with $a \in \mathbb{R}$. If $a > 0$,

$$P\psi(X - (\theta_0 + a)) = \int_{-\infty}^{\infty} p(x)\, \psi(x - (\theta_0 + a))\, dx$$

$$= \int_{-\infty}^{\infty} p(\theta_0 + (x - \theta_0))\, \psi((x - \theta_0) - a)\, dx$$

$$= \int_{-\infty}^{\infty} p(\theta_0 + y)\, \psi(y-a)\, dy$$

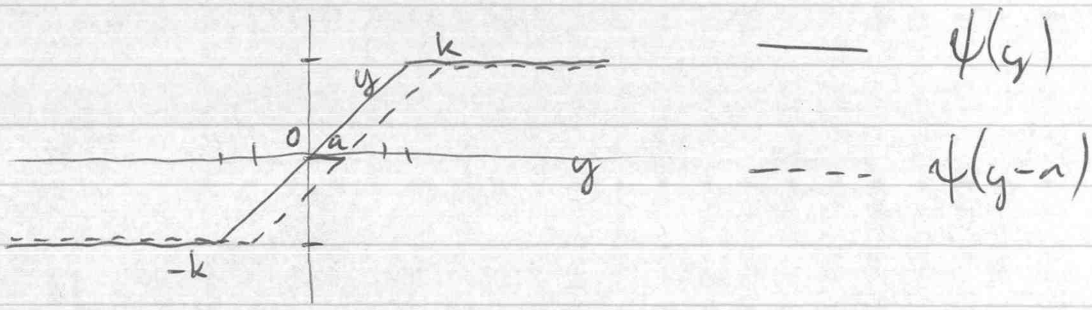
If $a=0$, then,

$$= \int_{-\infty}^{\infty} p(\theta_0+y)\, \psi(y)\, dy$$

$$= -k \int_{-\infty}^{-k} p(\theta_0+y)\, dy + k \int_{k}^{\infty} p(\theta_0+y)\, dy + \int_{-k}^{k} p(\theta_0+y)\, y\, dy = 0$$

the first two terms cancel and the third $= 0$.

If $a>0$, then $y-a < y \Rightarrow \psi(y-a) < \psi(y)$ $\left(-(k-a) \leq y \leq (k-a) \right)$



and $\psi(y-a) \leq \psi(y)$ everywhere else. Because $p(x)>0$, $P(-k+a \leq X-\theta_0 \leq k-a) > 0$ so for $a>0$

$\Psi(\theta_0+a) = P\psi(X-(\theta_0+a)) < P\psi(X-\theta_0) = 0$ and similarly, $\Psi(\theta_0-a)>0$. Use lemma 4.9 to conclude that $\hat{\theta}_n \xrightarrow{P} \theta_0$.

$\boxed{4.18}$ $\boxed{Q}$ Calculate the KLD between two exp. distr.

$\boxed{A}$ Consider P_λ and P_μ with Lebesgue densities

$$p_\lambda(x) = \lambda e^{-\lambda x}, \qquad p_\mu(x) = \mu e^{-\mu x}, \qquad (x \geq 0)$$

Then

$$\frac{p_\lambda}{p_\mu}(x) = \frac{\lambda}{\mu} e^{(\mu-\lambda)x}, \qquad (x \geq 0)$$

and

$$P_\lambda \log \frac{p_\lambda}{p_\mu} = \int_0^\infty \lambda e^{-\lambda x} \left((\mu-\lambda)x + \log \frac{\lambda}{\mu} \right) dx$$

$$= \left(\log \lambda - \log \mu\right) + (\mu-\lambda) \int_0^\infty \lambda x e^{-\lambda x} \, dx$$

$$= \left(\log \lambda - \log \mu\right) + \frac{\mu-\lambda}{\lambda} \int_0^\infty u e^{-u} \, du$$

$$= \log \lambda - \log \mu + \left(\frac{\mu}{\lambda} - 1\right) \left(-ue^{-u} \Big|_0^\infty + \int_0^\infty e^{-u} \, du \right)$$

$$= \log \frac{\lambda}{\mu} + \left(\frac{\mu}{\lambda} - 1\right)$$

$$P_\theta \dot{l}_\theta \dot{l}_\theta = P_\theta\left(-\frac{X}{\theta} + \frac{1}{\theta^2}\right)\left(-\frac{X}{\theta} + \frac{1}{\theta^2}\right)$$

$$= P_\theta\left(\frac{X^2}{\theta^2} - 2\frac{X}{\theta^3} + \frac{1}{\theta^4}\right)$$

$$= P_\theta\left(\frac{X^2}{\theta^2}\right) - \frac{1}{\theta^4} = \frac{1}{\theta^2}\left(P_\theta X^2 - (P_\theta X)^2\right)$$

$$= \frac{1}{\theta^2} \mathrm{Var}_\theta(X) = \frac{1}{\theta^3}$$