Assignment-1

If $X_1, \ldots X_n$ is a seq of Random variables.

(1)
(a) $X_n \xrightarrow{a.s} X$ means that

$$Pr\left(\lim_{n\to\infty} X_n = X\right) = 1$$

(b) $X_n \xrightarrow{P} X$ means that

$$Pr\left(\lim_{n\to\infty} |X_n - X| \geq \epsilon\right) = 0 \quad \text{for any } \epsilon > 0$$

(c) $X_n \xrightarrow{MS} X$ means that

when $\lim_{n\to\infty} E\left[(X_n - X)^2\right] = 0$

(d) $X_n \xrightarrow{d} X$

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

where F_{X_n} and F_X are cumulative distribution functions of random variables X_n and X respectively.

(2) $\{X_n ; (n \geq 1)\}$ $\quad X_n = \cos[n\theta]$

(a) The distribution of X_n is the same for all n, so the sequence converges in distribution to any random variable with distribution X_1

To check Mean square convergence, use the fact $\cos(a)\cos(b) = \left(\cos(a+b) + \cos(a-b)\right)/2$ to calculate that $E[X_n X_m] = 1/2$ if $n = m$ and $E[X_n X_m] = 0$ if $n \neq m$. Therefore, $\lim_{n\to\infty} E(X_n X_m)$ doesnot exist so the sequence (X_n) doesnot satisfy the cauchy criterion for m.s convergence so doesn't converge in m.s sense. Since it is a bounded sequence, it therefore doesnot converge in probability (Bounded sequences, convergence in P. implies convergence m.s).

Another way to show it is as follows, Distribution for $X_n - X_{2n}$ is the same for all n, so the sequence doesn't satisfy the cauchy criterion for convergence in probability.

Thus only $X_n \xrightarrow{d} X$ satisfies.

(b) $\{Y_n : n \geq 1\}$ defined by $Y_n = \left[\left(1 - \frac{\theta}{\pi}\right)\right]^n$

If θ is such that $0 < \theta < 2\pi$ then

$0 < \left|1 - \frac{\theta}{\pi}\right| < 1$ so that $\lim_{n \to \infty} Y_n = 0$

for such θ

since $P\{0 < \theta < 2\pi\} = 1$ it follows that

Y_n converges to 0 in the almost sure (a.s) sense and hence also in the 'p' and d senses.

Since the sequence is bounded, it also converges to zero in the meansquared (m.s) sense.

thus all four 'a.s', 'p', 'd', 'm.s' satisfies.
(1) (2) (3) (4)

(6)

(3) Moment generating function of $X \sim N(0, \sigma^2)$

$$E(e^{sx}) = \int_{-\infty}^{\infty} \frac{e^{sx} \cdot e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi}\,\sigma} dx$$

$$= \int_{-\infty}^{\infty} \frac{e^{-\left(\frac{x^2 - 2\sigma^2 s x}{2\sigma^2}\right)}}{\sqrt{2\pi}\,\sigma} dx$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}\left[x^2 - 2\sigma^2 s x + \sigma^4 s^2\right] + \frac{\sigma^4 s^2}{2\sigma^2}} dx$$

$$= \frac{e^{\frac{\sigma^2 s^2}{2}}}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{(x - \sigma^2 s)^2}{2\sigma^2}} dx$$

$$= \left[e^{\sigma^2 s^2/2}\right] \left[\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{(x-\sigma^2 s)^2}{2\sigma^2}} dx\right]$$

$$= e^{\sigma^2 s^2/2}$$

(4) (a) Chebyshev inequality

$$P_r\left(|X-\mu| > \epsilon\right) \leq \frac{\sigma^2}{\epsilon^2}$$

where $\sigma = E((X-\mu)^2)$

Proof :-

Let $Y = (x-\mu)^2$ be a R.V which is nonnegative

Now

$Pr(|x-\mu| \geq \epsilon) = Pr((X-\mu)^2 > \epsilon^2)$

this is true. $\phi(x) = x^2$ is non negative and strictly monotonic and increasing.

$$Pr(|x-\mu| \geq \epsilon) = Pr((|X-\mu|)^2 \geq \epsilon^2) \leq \frac{E((X-\mu)^2)}{\epsilon^2}$$

This is from Markov inequality

$$\boxed{Pr(Y \geq \epsilon^2) \leq \frac{E(Y)}{\epsilon^2} \quad \text{for } \epsilon > 0}$$

(b) Chernoff bound

Suppose $t > 0$ & X be any Random variable

$$Pr(X \geq \epsilon) = Pr(Xt \geq \epsilon t) = Pr(e^{Xt} \geq e^{\epsilon t})$$

Since $\phi(x) = e^{t(x)}$ for $t > 0$ is a monotonic and increasing function.

Now from Markov inequality since $Y = e^{Xt}$ is a non negative random variable

$$Pr(e^{Xt} \geq e^{\epsilon t}) \leq \frac{E(e^{Xt})}{e^{\epsilon t}} \quad \text{Since for } Y = e^{xt}$$

$$\boxed{Pr(Y \geq e^{\epsilon t}) \leq \frac{E(Y)}{e^{\epsilon t}}} \quad \text{for all } e^{\epsilon t} > 0$$

(5)

(a) $X \sim \{0, 1\}$

$Pr(X=0) = 1 - \frac{1}{k}$, $Pr(X=1) = 1/k$,

$E(X) = \frac{1}{k}$

$Pr(X \geq kE(X)) \leq \frac{E(X)}{kE(X)}$

$$\boxed{Pr(X \geq k(1/k)) \leq \frac{1}{k}}$$

So Here

$Pr(X \geq 1) \leq \frac{1}{k}$ but equality holds here

$\boxed{Pr(X \geq 1) = \frac{1}{k}}$ thus tighter bound in this case

(5)

(6) Say $X \sim N(0,1)$

$Pr(|X| > \epsilon) \leq \frac{1}{\epsilon^2}$ (Chebyshev) — (a)

also,

$Pr(e^{Xt} > e^{\epsilon t}) \leq e^{\frac{\sigma^2 t^2 - 2\epsilon t}{2}}$ (Chernoff) — (b)

Now we can see $\forall t > 0$
RHS (a), (b)

$e^{(t^2 - 2\epsilon t)/2} \leq \frac{1}{2\epsilon^2}$

$$\boxed{e^{t^2/2} \leq \frac{e^{\epsilon t}}{2\epsilon^2}}$$

which is true for $\boxed{\epsilon < 0}$.

$N(0,1)$ is a case where chernoff is tighter than Chebyshev. $\boxed{\text{for } \epsilon < 0}$

Scanned by CamScanner

(5) (c)

$S_{100} = \sum_{i=1}^{100} X_i$ where X_i's are i.i.d r.v.

X_i are non negative

Chebyshev inequality

$P(|(S_{100} - \mu)| > x) \leq \frac{Var(S_{100})}{\epsilon^2}$ Hence S_{100}

$\mu = 2750 \quad (27.50 \times 100)$

$var = 16875 \; (168.75 \times 100)$

$\sigma = 250$

$\frac{1}{2}\left(P(S_{100} > 3000)\right) = \frac{1}{2}\left(P(|S_{100} - 2750| > 250)\right) \leq \frac{16875}{2(250)^2}$

$\leq \frac{0.27}{2}$

≤ 0.135

CLT

$P(S_{100} > 3000) = P\left(\frac{S_{100}}{100} > 30\right)$

$\left(\frac{S_{100}}{100}\right) \sim N\left(\mu, \frac{\sigma^2}{100}\right) = X \text{(say)}$

$Pr(X > 30) = Pr\left(\left(\frac{X-\mu}{\sigma}\right) > \frac{30-\mu}{\sigma}\right)$

$= \phi\left(\left(\frac{30-27.5}{13}\right) \times 10\right)$

$= \phi(1.923) = 0.027.$

Scanned by CamScanner

(6) Let $X_1, \ldots X_N$ be iid random variables with each X_i having mean μ, variance σ^2.

Suppose,

$$\boxed{S = \frac{\sum_{i=1}^{n} X_i}{N}}$$

$$E(S) = \mu, \ \mathrm{var}(S) = E\left((S-\mu)^2\right) = \frac{\sigma^2}{n}$$

From chebyshev's inequality

$$Pr(|S - \mu| \geq \epsilon) \leq \frac{\sigma^2}{n\epsilon^2}$$

$$\lim_{n\to\infty} Pr\left(\left|\frac{\sum_{i=1}^{N} X_i}{N} - \mu\right| \geq \epsilon\right) \leq \lim_{n\to\infty} \frac{\sigma^2}{n\epsilon^2}$$

$$Pr\left(\lim_{n\to\infty} \left|\frac{\sum_{i=1}^{N} X_i}{N} - \mu\right| \geq \epsilon\right) \leq \lim_{n\to\infty} \frac{\sigma^2}{n\epsilon^2}$$

(probabability cannot be less than 0)

$$\boxed{Pr\left(\lim_{n\to\infty} \left|\frac{\sum_{i=1}^{N} X_i}{N} - \mu\right| \geq \epsilon\right) \to 0}$$

Thus proving weak Law of Large numbers,

$$\boxed{Pr\left(\lim_{n\to\infty} \left|\frac{\sum_{i=1}^{N} x_i}{N} - \mu\right| \geq \epsilon\right) = 0}$$

7)

(a) Rademacher Random variable

$$X = \begin{cases} 1 & P(X=1) = 1/2 \\ -1 & P(X=-1) = 1/2 \end{cases}$$

$$E(\exp(tX)) = P(x=-1)e^{-t} + P(X=1)e^{t}$$

$$= \frac{e^{-t}+e^{t}}{2} = \cosh(t) = \sum_{n\geq 0} \frac{t^{2n}}{(2n)!}$$

$$E(\exp(tX)) = \sum_{n\geq 0} \frac{t^{2n}}{(2n)!} \leq \sum_{n\geq 0} \frac{(t^2)^n}{2^n n!} \quad \because 2^n n! \leq (2n)! \ \forall n.$$

Now we know $\sum_{n\geq 0} \frac{(t^2)^n}{2^n n!} = \exp(t^2/2)$

$$\boxed{E(e^{tX}) \leq e^{t^2/2}}$$

Thus X → Rademacher random variable is 1-sub gaussian.

(b). A gaussian random variable with mean μ and variance σ^2.

$X \sim N(0, \sigma^2)$

$$E(e^{tX}) = \int_{-\infty}^{\infty} \frac{e^{tx} \, e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi}\,\sigma} dx$$

after rearranging as in eqn (3)

$$= \frac{e^{\sigma^2 t^2/2}}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} e^{-(x-\sigma^2 s)^2/2\sigma^2} dx$$

$$= e^{\sigma^2 t^2/2} \left[\frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} e^{-\frac{(x-\sigma^2 s)^2}{2\sigma^2}} dx \right]$$

$$= \underline{\underline{e^{\frac{\sigma^2 t^2}{2}}}}$$

Thus is σ^2-sub-gaussian

(c) A random variable that is zero mean and bounded in the interval $[a, b]$

from Hoeffding's Lemma

The random variable X with $E(X) = 0$ and bounded $a \le x \le b$ then for any $s \in \mathbb{R}$

$$\mathbb{E}\left(e^{sX}\right) \leq e^{\frac{s^2(b-a)^2}{8}}$$

on comparison

$$\sigma^2 = \frac{(b-a)^2}{4}$$

ie. X is $\frac{(b-a)^2}{4}$ - sub gaussian

(8) To prove

$$P(|X-\mu| \geq t) \leq e^{\left(\frac{-t^2}{2\sigma^2}\right)} \quad \forall\, t>0$$

Now

given X is a σ^2 sub gaussian with mean μ

which means

$$E\left(e^{\lambda(X-\mu)}\right) \leq e^{\frac{\sigma^2\lambda^2}{2}} \quad -(1) \quad \forall\, \lambda \in \mathbb{R}.$$

→ (1) we first prove

$$P((X-\mu) \geq t) = P\left(e^{(X-\mu)s} \geq e^{st}\right) \quad \forall\, s>0$$

From chernoff bound $P\left(e^{(X-\mu)s} \geq e^{st}\right) \leq e^{-st}\left[E\left(e^{(X-\mu)s}\right)\right]$

$$P((X-\mu) \geq t) \leq e^{-st}\, E\left(e^{(X-\mu)s}\right)$$

Now from sub gaussian definition

$$\boxed{E\left(e^{(X-\mu)s}\right) \leq e^{\frac{\sigma^2 s^2}{2}}} \quad \forall\, s \in \mathbb{R}$$

substitute it back.

$$P((X-\mu) \geq t) \leq e^{-st}\, e^{\sigma^2 s^2/2}$$

$$P((X-\mu) \geq t) \leq e^{\left(-st + \frac{\sigma^2 s^2}{2}\right)} \quad \forall\, s>0$$

to find a tighter bound we have to find value of s for which $\phi(s) = e^{-st + \frac{\sigma^2 s^2}{2}}$ is minimum.

Since $\phi(x) = e^x$ is increasing and monotonic.

$f(s) = (-st) + \frac{\sigma^2 s^2}{2}$

$f'(s) = -t + \frac{2\sigma^2 s}{2} = 0$

$\Rightarrow \boxed{s = \frac{+t}{\sigma^2}}$ substitute it back we get minimum values of $e^{-st + \frac{\sigma^2 s^2}{2}}$ which is $e^{\frac{-t^2}{2\sigma^2}}$

$P((X-\mu) \geq t) \leq e^{\left(-\frac{t^2}{\sigma^2} + \frac{\sigma^2 t^2}{2\sigma^4}\right)}$

$$\boxed{P((X-\mu) \geq t) \leq e^{-\frac{t^2}{2\sigma^2}} \quad \forall t > 0}$$ — (a)

$\rightarrow$ (2)

Now showing

$P((X-\mu) \leq -t) = P((\mu - X) \geq t)$

Since $\phi(x) = e^x$ is increasing (strictly) function $\boxed{\text{for } s > 0}$

$= P(e^{s(\mu - x)} \geq e^{st})$

now by applying chernoff bound,

$P((X-\mu) \leq -t) = P(e^{s(\mu-x)} \geq e^{st}) \leq e^{-st} E(e^{s(\mu-x)})$

If x is σ^2 sub gaussian with symmetry $(-x)$ is also σ^2 sub gaussian

$P((X-\mu) \leq -t) \leq e^{-st + \frac{\sigma^2 s^2}{2}}$

$$P((X-\mu) \leq -t) \leq e^{-St+\frac{\sigma^2 S^2}{2}} \quad \forall S \geq 0$$

As proved above for a tighter bound we take the minimum of the function value $e^{-St+\frac{\sigma^2 S^2}{2}}$ happens at $S = \frac{t}{\sigma^2}$

and

$$\boxed{P((X-\mu) \leq -t) \leq e^{-\frac{t^2}{2\sigma^2}} \quad \forall t > 0} \quad (b)$$

NOW:

$$P(|X-\mu| \geq t) = P(((X-\mu) \geq t) \cup ((X-\mu) \leq -t))$$

$$P(|X-\mu| \geq t) \leq P((X-\mu) \geq t) + P((X-\mu) \leq -t)$$

from (a) & (b)

$$P(|X-\mu| \geq t) \leq e^{-t^2/2\sigma^2} + e^{-t^2/2\sigma^2}$$

$$\boxed{P(|X-\mu| \geq t) \leq 2e^{-\frac{t^2}{2\sigma^2}} \quad \forall t > 0}$$

e (1) X_i's are independent mean zero, σ_i^2 subgaussian sub gaussian random variables then

Let

$Z = \sum_{i=1}^{n} X_i$ be a random variable

Now

$$E(e^{tZ}) = E\left(e^{t\left(\sum_{i=1}^{n} X_i\right)}\right)$$

Now we know that for independent random variables

$$pdf(x_1, x_2, \cdots x_n) = pdf(x_1) \cdot pdf(x_2) \cdots pdf(x_n)$$

Hence

$$E(e^{tZ}) = E\left(e^{t\left(\sum_{i=1}^{n} X_i\right)}\right) = E(e^{tX_1}) \cdot E(e^{tX_2}) \cdots E(e^{tX_n}) \quad -(1)$$

Since we know that,

$$E(e^{tX_i}) \leq e^{\left(t^2\sigma_i^2/2\right)} \quad \text{for all } t \in \mathbb{R}.$$

Thus

$$E(e^{tX_1}) \cdot E(e^{tX_2}) \cdots E(e^{tX_n}) \leq e^{\frac{t^2}{2}\left(\sum_{i=1}^{n}\sigma_i^2\right)} \quad -(2)$$

Sub (2) in (1)

$$E(e^{tZ}) = E\left(e^{t\left(\sum_{i=1}^{n} X_i\right)}\right) \leq e^{\frac{t^2}{2}\left(\sum_{i=1}^{n}\sigma_i^2\right)}$$

Thus $Z = \sum_{i=1}^{n} X_i$ is $\sum_{i=1}^{n}\sigma_i^2$ - sub gaussian.

Scanned by CamScanner

(10) Hoeffdings inequality

If $S_n = \sum_{i=1}^{n} X_i$ where X_i's are independent RVs with $a_i \leq X_i \leq b_i$ then,

$$P\left(\frac{S_n}{n} - E\left(\frac{S_n}{n}\right) \geq t\right) \leq e^{-st} E\left(e^{s\left(\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right)}\right) \quad (1)$$

$\boxed{t > 0}$

we show (1) as

Let $Z = \frac{S_n}{n} - E\left(\frac{S_n}{n}\right)$

$\boxed{E(Z) = 0}$ and since X_i is bounded between $a_i \leq X_i \leq b_i$ thus Z is also bounded.

Now

$$P(Z \geq t) = P(e^{sZ} \geq e^{ts})$$

Since $e^{s(x)}$ for $s > 0$ is monotonic and increasing function $\boxed{t > 0}$

Now by chernoff bound

$$P(Z \geq t) = P(e^{sZ} \geq e^{ts}) \leq e^{-ts} E_Z(e^{sZ})$$

$\boxed{\text{Chernoff bound}}$

Now,

$$P\left(\frac{S_n}{n} - E\left(\frac{S_n}{n}\right) \geq t\right) \leq e^{-st} E\left(e^{s\left(\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right)}\right)$$

$$\leq e^{-st} \prod_{i=1}^{n} E\left(e^{s\frac{[x_i - E(x_i)]}{n}}\right) \quad -(3)$$

Since all x_i's are independent.

Taking $E\left(e^{s\left(\frac{x_i - E(x_i)}{n}\right)}\right)$

$\rightarrow$ Now $Y = X_i - E(x_i)$ has $E(Y) = 0$ and is bounded

$a_i \leq x_i \leq b_i \quad \Rightarrow \left(\frac{a_i - E(x_i)}{n}\right) \leq \left(\frac{X_i - E(x_i)}{n}\right) \leq \left(\frac{b_i - E(x_i)}{n}\right)$

thus we can apply hoeffding's Lemma:

$$E\left(e^{s\frac{(x_i - E(x_i))}{n}}\right) \leq e^{s^2(b_i - a_i)^2/8n^2} \quad -(4)$$

$\forall s > 0$

Substitute (4) in (3).

$$P\left(\left(\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right) \geq t\right) \leq e^{-st}\left(\prod_{i=1}^{n} e^{s^2(b_i - a_i)^2/8n^2}\right)$$

$$\leq e^{-st}\left(e^{\frac{s^2}{8n^2}\left(\sum_{i=1}^{n}(b_i - a_i)^2\right)}\right)$$

$$\leq \exp\left(-st + \frac{s^2}{8n^2}\left(\sum_{i=1}^{n}(b_i - a_i)^2\right)\right)$$

$\forall s > 0$

we have to minimize $f(s) = -st + \frac{s^2}{8n^2}\left(\sum_{i=1}^{n}(b_i - a_i)^2\right)$

with respect to s because $\phi(x) = e^x$ is monotonic and increasing fn.

$f'(s) = 0$

$$\Rightarrow -t + \frac{s}{4n^2}\left(\sum_{i=1}^{n}(b_i - a_i)^2\right) = 0$$

$$\boxed{s = \frac{4tn^2}{\sum_{i=1}^{n}(b_i-a_i)^2}} \quad -(5)$$

sub (5) back to find a tighter bound.

$$P\left(\frac{S_n}{n} - E\left(\frac{S_n}{n}\right) \geq t\right) \leq \exp\left(-\frac{4t^2n^2}{\sum_{i=1}^{n}(b_i-a_i)^2} + \frac{16}{8}\left(\frac{t^2n^4}{n^2}\right)\frac{1}{\sum_{i=1}^{n}(b_i-a_i)^2}\right)$$

$$\leq \exp\left(-\frac{4t^2n^2}{\sum_{i=1}^{n}(b_i-a_i)^2} + \frac{2t^2n^2}{\sum_{i=1}^{n}(b_i-a_i)^2}\right)$$

$$\leq \exp\left(\frac{-2t^2n^2}{\sum_{i=1}^{n}(b_i-a_i)^2}\right)$$

$$\boxed{P\left(\frac{S_n}{n} - E\left(\frac{S_n}{n}\right) \geq t\right) \leq \exp\left(-\frac{2t^2n^2}{\sum_{i=1}^{n}(b_i-a_i)^2}\right)}$$

$\forall\ t > 0$

ie.

$$P\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - E(X_i)) \geq t\right) \leq \exp\left(-\frac{2t^2n^2}{\sum_{i=1}^{n}(b_i-a_i)^2}\right)$$

$\forall\ t > 0$

(10) For the other part,

$$y_i = \frac{E(x_i) - x_i}{n}$$

$$P\left(\sum_{i=1}^{n} y_i \geq t\right) = P\left(\frac{1}{n}\left[\sum_{i=1}^{n} x_i - E(x_i)\right] \leq -t\right) \quad \text{— (a)} \quad \forall\, t > 0$$

Now applying Hoeffdings Lemma to y_i

Since $E(y_i) = 0$ and

$$\frac{E(x_i) - b_i}{n} \leq y_i \leq \frac{E(x_i) - a_i}{n} \quad \text{ie } y_i \text{ is bounded.}$$

By Hoeffdings Lemma

$$E(e^{\lambda y_i}) \leq e^{\frac{\lambda^2 (E(x_i) - b_i - E(x_i) + a_i)^2}{8n^2}} \quad \forall\, \lambda > 0$$

$$E(e^{\lambda y_i}) \leq e^{\frac{\lambda^2 (b_i - a_i)^2}{8n^2}} \quad \text{— } (*)$$

Now from (a)

$$P\left(\frac{1}{n}\left[\sum_{i=1}^{n} x_i - E(x_i)\right] \leq -t\right) = P\left(\sum_{i=1}^{n} y_i \geq t\right) \quad \forall\, t > 0$$

Now since $f(x) = e^x$ is strictly increasing (monotone)

$$= P\left(e^{\lambda \sum_{i=1}^{N} y_i} \geq e^{\lambda t}\right) \quad \forall\, t > 0$$

from chernoff bound;

$$P\left(\sum_{i=1}^{n} y_i \geq t\right) = P\left(e^{\lambda \sum_{i=1}^{n} y_i} \geq e^{\lambda t}\right) \leq e^{-\lambda t} E\left(e^{\lambda \sum_{i=1}^{n} y_i}\right) \quad \boxed{\forall \lambda > 0}$$

$$P\left(\sum_{i=1}^{n} y_i \geq t\right) \leq e^{-\lambda t} E\left(e^{\lambda \sum_{i=1}^{n} y_i}\right)$$

Since each y_i's (x_i's) are independent

$$P\left(\sum_{i=1}^{n} y_i \geq t\right) \leq e^{-\lambda t} \prod_{i=1}^{n} E\left(e^{\lambda y_i}\right)$$

From hoeffding's Lemma in (*)

$$P\left(\frac{1}{n}\sum_{i=1}^{n} [E(x_i) - x_i] \geq t\right) \leq e^{-\lambda t}\left[\prod_{i=1}^{n} e^{\frac{\lambda^2 (b_i - a_i)^2}{8n^2}}\right]$$

$$\leq e^{-\lambda t + \frac{\lambda^2 \sum_{i=1}^{n} (b_i - a_i)^2}{8n^2}} \quad \boxed{\forall \lambda > 0}$$

Now the tighter bound is obtained at the minimum value of $f(\lambda) = e^{-\lambda t + \lambda^2 \left(\sum_{i=1}^{n} (b_i - a_i)^2\right)/8n^2}$.

$$f'(\lambda) = -t + \frac{2\lambda}{8n^2}\left(\sum_{i=1}^{n} (b_i - a_i)^2\right) = 0$$

$$\boxed{\lambda = \frac{4n^2 t}{\sum_{i=1}^{n} (b_i - a_i)^2}} \rightarrow \text{substituting this back.}$$

$$P\left(\frac{1}{n}\sum_{i=1}^{n} [E(x_i) - x_i] \geq t\right) \leq e^{\left[-\frac{2n^2 t^2}{\sum_{i=1}^{n} (b_i - a_i)^2}\right]}$$

$$\Rightarrow P\left[\frac{1}{n}\sum_{i=1}^{n} [x_i - E(x_i)] \leq -t\right] \leq e^{-\frac{2n^2 t^2}{\sum_{i=1}^{n} (b_i - a_i)^2}} \quad \boxed{\forall t > 0}$$

(11)

(a) To prove The exponential random variable with parameter λ is sub exponential

$$P_X(x) = \lambda e^{-\lambda x}$$

$$E(X) = \frac{1}{\lambda}$$

$$E\left(e^{s(X-\frac{1}{\lambda})}\right) = \lambda e^{\frac{-s}{\lambda}} \int_0^\infty e^{s(x)} e^{(-\lambda x)} dx$$

$$= \lambda e^{-s/\lambda} \int_0^\infty e^{(s-\lambda)x} dx$$

finite if $s < \lambda$

$$E\left[e^{s(X-\frac{1}{\lambda})}\right) = \frac{\lambda}{\lambda - s} e^{-s/\lambda}$$

$$= \frac{1}{1-\left(\frac{s}{\lambda}\right)} e^{-s/\lambda}$$

applying taylor series and analysing

$$E\left(e^{s(X-\frac{1}{\lambda})}\right) = \frac{e^{-(s/\lambda)}}{1-\left(\frac{s}{\lambda}\right)} \leq e^{s^2/\lambda^2} \quad \forall |\lambda| < \frac{1}{2}$$

thus it is sub exponential with

$$v = \sqrt{2}/\lambda, \; b = \frac{2}{\lambda}$$

(b) $X \sim \chi^2_1$ $\rightarrow$ chi square distribution with 1 degree of freedom

Now

$X = Z^2$ where $Z \sim N(0,1)$

The distribution of X is chi square distribution with 1 degree of freedom

$$\text{Pdf}(z) = \frac{1}{\sqrt{2\pi}} e^{z^2/2}$$

Now the mean (X) is $\underline{E(Z^2) = 1}$

Thus

$E(\exp(\lambda(z^2-1)))$

$$E(\exp(\lambda(X-1))) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(\lambda(z^2-1)) e^{-\frac{z^2}{2}} dz$$

$$= e^{-\lambda} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{z^2\left[\frac{2\lambda - 1}{2}\right]} dz \right)$$

On solving

$$= \frac{e^{-\lambda}}{\sqrt{1-2\lambda}}$$

$$MGF(x) = \frac{1}{\sqrt{1-2\lambda}} e^{-\lambda} \quad \forall\, |\lambda| < 1/2.$$

Now we know that

$$\frac{e^{-\lambda}}{\sqrt{1-2\lambda}} \leq e^{2\lambda^2} \quad \forall\, |\lambda| < \frac{1}{4}$$

This can be proved by taking taylor expansion on LHS.

$$\text{thus} \quad E(e^{\lambda(X-1)}) \leq e^{2\lambda^2} \quad \forall\, |\lambda| < \frac{1}{4}$$

$$E(e^{\lambda(X-1)}) \leq e^{\frac{4\lambda^2}{2}} \quad \forall\, |\lambda| < \frac{1}{4}$$

on comparison with $e^{\frac{v^2\lambda^2}{2}}$

$$\boxed{v = 2 \quad , b = 4}$$

Thus chi square distribution with 1 degree of freedom is sub exponential with parameter (2,4)

In general,
Now chi-square distribution with n degree of freedom is sub exponential with parameters

$(2n, 4)$

Since

$Y \sim \chi^2$ with n degree of freedom is

$$Y = \sum_{i=1}^{N} Z_i^2$$ where Z_i's are independent

Thus from the property

$Y = \sum_{i=1}^{N} Z_i^2$ is sub exponential with parameters

$$\left(\sum_{i=1}^{n} v_i, \max_i b_i \right) = (2n, 4).$$

To prove :-

Suppose X is a (v, b) sub exponential random variable with mean μ, derive the tail bound $P(X \geq \mu + t)$ $\forall t > 0$.

Proof :-

$$P(X \geq \mu + t) = P((X - \mu) \geq t)$$

Now since $\phi(x) = e^{x}$ is monotonically increasing function

$$= P(e^{(X-\mu)\lambda} \geq e^{\lambda t}) \quad \boxed{\lambda > 0}$$

Now applying chernoff bound;

$$P(X \geq \mu + t) = P(e^{(X-\mu)\lambda} \geq e^{\lambda t}) \leq e^{-\lambda t} E(e^{(X-\mu)\lambda})$$

Now for a sub exponential random variable

$$E[e^{\lambda(X-\mu)}] \leq e^{\frac{v^2 \lambda^2}{2}} \quad \forall |\lambda| < \frac{1}{b}$$

where (v, b) are non negative parameters

$$P(X \geq \mu + t) \leq e^{-\lambda t + \frac{v^2 \lambda^2}{2}} \quad \forall \lambda \in [0, \tfrac{1}{b})$$

Scanned by CamScanner

To find tighter bounds we have to minimize

$e^{-\lambda t + v^2\lambda^2/2}$ wrt. (λ)

Since $f(x) = e^x$ is stricly increasing fn.

we have to minimize $f(\lambda) = -\lambda t + \frac{v^2\lambda^2}{2}$ wrt λ.

$f'(\lambda) = -t + v^2\lambda = 0$ $\quad f''(\lambda) = v^2 > 0$

$\boxed{\lambda = \frac{t}{v^2}}$

thus $\lambda = \frac{t}{v^2}$ is minimum

Now there arises 2 conditions if

① → $\lambda = \frac{t}{v^2}$ is in between 0 and $\frac{1}{b}$

$\boxed{t > 0, v > 0, b > 0}$

or ② → $\lambda = \frac{t}{v^2} > \frac{1}{b}$

① solving first condition, if

→ $(\lambda) = \frac{t}{v^2}$ is in between 0 and $\frac{1}{b}$

$0 \quad t/v^2 \quad 1/b$

meaning $\frac{t}{v^2} \leq \frac{1}{b} \Rightarrow t \leq \frac{v^2}{b}$

For $0 \leq t \leq \frac{v^2}{b}$.

$e^{-\lambda t + \frac{v^2\lambda^2}{2}}$ attains minima at $\lambda = \frac{t}{v^2}$ since $\frac{t}{v^2} \leq \frac{1}{b}$

thus

$$P(X \geq \mu + t) \leq e^{-\frac{t^2}{2v^2}} \qquad 0 \leq t \leq \frac{v^2}{b}$$

(2)

condition (2) if $t > \frac{v^2}{6}$ then

$\lambda = \frac{t}{v^2} > \frac{1}{6}$ hence not lying in $\left[0, \frac{1}{6}\right]$

Thus by observing the fn

$\phi(\lambda) = e^{-\lambda t + v^2\lambda^2/2}$ which is monotonically decreasing in $\left[0, \frac{t}{v^2}\right]$ which contains the set $\left[0, \frac{1}{6}\right]$

since $\boxed{\frac{1}{6} \le \frac{t}{v^2}}$

The local minima for this set happens to be at $\boxed{\lambda = \frac{1}{6}}$.

Thus substituting it back

$P(X \geq \mu + t) \leq e^{\left(-t/6 + \frac{v^2}{2 \cdot 6^2}\right)}$ — (a)

Now we know since $\boxed{t \geq \frac{v^2}{6}}$

thus $e^{t} \geq e^{v^2/6} \Rightarrow e^{\frac{t}{2 \cdot 6}} \geq e^{\frac{v^2}{2 \cdot 6^2}}$ since $6 > 0$

$\Rightarrow e^{-\frac{t}{6}} e^{\frac{t}{2 \cdot 6}} \geq e^{-\frac{t}{6} + \frac{v^2}{2 \cdot 6^2}}$

from (b) $e^{\left(-\frac{t}{6} + \frac{v^2}{2 \cdot 6^2}\right)} \leq e^{\left(-\frac{t}{6} + \frac{t}{2 \cdot 6}\right)}$ — (b)

$$\boxed{P(X \geq \mu + t) \leq e^{-\frac{t}{2 \cdot 6}} \text{ for } t > \frac{v^2}{6}}$$

Thus

$$P(X \geq \mu + t) \leq \begin{cases} e^{-\frac{t^2}{2\nu^2}} & \text{if } 0 \leq t \leq \nu^2/b \\ e^{-\frac{t}{2b}} & \text{if } t > \frac{\nu^2}{b} \end{cases}$$

(13) For independent x_i, sub exponential with parameters (ν_i, b_i) then, let the sum

$$X = \sum_{i=1}^{N} x_i ,$$

$$E(e^{\lambda(X)}) = E(e^{\lambda(\sum_{i=1}^{N} x_i)})$$

Since all x_i's $\forall i = 1, \dots N$ are independent

$$E(e^{\lambda(X)}) = E(e^{\lambda x_1}) \cdot E(e^{\lambda x_2}) \dots E(e^{\lambda x_N})$$

Now

Let μ_i be the mean of X_i

then the mean of $X = \sum_{i=1}^{N} x_i$ is $\mu = \sum_{i=1}^{N} \mu_i$

Thus,

$$E(e^{\lambda(X - \mu)}) = E(e^{\lambda(\sum_{i=1}^{N} x_i - \sum_{i=1}^{N} \mu_i)})$$

$$E\left(e^{\lambda(X-\mu)}\right) = E\left(e^{\lambda \sum_{i=1} (X_i - \mu_i)}\right)$$

Since all X_i's are independent $\forall i \in (1, \cdots N)$

Thus

$$E\left(e^{\lambda(X-\mu)}\right) = E\left(e^{\lambda(X_1 - \mu_1)}\right) \cdots\cdots E\left(e^{\lambda(X_N - \mu_N)}\right) \quad \leftarrow (a)$$

Since each X_i is sub exponential with parameters (v_i, b_i)

$$E\left(e^{\lambda(X_i - \mu_i)}\right) \le e^{v_i^2 \lambda^2 / 2} \quad \forall |\lambda| < \frac{1}{b_i}.$$

For the bound to be valid for all X_i's the region of $|\lambda|$ should be less than $\min\left(\left(\frac{1}{b_1}\right), \left(\frac{1}{b_2}\right), \ldots \left(\frac{1}{b_n}\right)\right)$

$$\boxed{|\lambda| < \frac{1}{\left[\max_i (b_i)\right]}}$$

ie intersection will be valid for $|\lambda| < \frac{1}{\left[\max_i b_i\right]}$

From (a)

$$E\left[e^{\lambda(X-\mu)}\right] = E\left(e^{\lambda(X_1-\mu_1)}\right) \cdots E\left(e^{\lambda(X_N - \mu_N)}\right)$$

$$\le e^{\frac{\lambda^2}{2} \sum_{i=1}^{N} v_i^2} \quad \text{for } |\lambda| < \frac{1}{\max_i b_i}$$

Thus

$$E\left(e^{\lambda\left[\sum_{i=1}^{N} X_i - \mu\right]}\right) \le e^{\frac{\lambda^2}{2}\left[\sum_{i=1}^{N} v_i^2\right]} \le e^{\frac{\lambda^2}{2}\left(\sum v_i\right)^2}$$

Thus $\sum_{i=1}^{N} X_i$ is sub exponential with

Parameters $\left(\sum_{i=1}^{N} v_i,\ b_*\right)$ where $b_* = \max_i b_i$

$E\left(e^{\lambda(\sum x_i - \sum \mu_i)}\right) \leq e^{\lambda^2 (\sum v_i)^2 / 2}$ $\forall |\lambda| < \frac{1}{\max_i b_i}$

Property used

$$\prod_{i=1}^{n} e^{v_i^2 \lambda^2 / 2} = e^{\frac{\lambda^2 (\sum_i v_i^2)}{2}} \leq e^{\frac{\lambda^2 (\sum v_i)^2}{2}}$$

Scanned by CamScanner

(14)

If X is a random variable with mean μ and variance σ^2 that satisfies

$$E[(x-\mu)^k] \leq \frac{1}{2} k! \sigma^2 b^{k-2} \quad (1)$$ then

it is said to satisfy bernstein bound

$$E(\exp(\lambda(x-\mu))) = E\left[1 + \frac{\lambda(x-\mu)}{1!} + \frac{\lambda^2(x-\mu)^2}{2!} + \sum_{k=3}^{\infty} \lambda^k \frac{E((x-\mu)^k)}{k!}\right]$$

$$= 1 + \frac{\lambda E(x-\mu)}{1!} + \frac{\lambda^2 E[(x-\mu)^2]}{2!} + \sum_{k=3}^{\infty} \frac{\lambda^k E(x-\mu)^k}{k!}$$

Now $E(x-\mu) = 0$, $E((x-\mu)^2) = \sigma^2$

$$E(\exp(\lambda(x-\mu))) = 1 + \frac{\lambda^2\sigma^2}{2!} + \sum_{k=3}^{\infty} \frac{\lambda^k E((x-\mu)^k)}{k!}$$

From (1) $\boxed{E[(x-\mu)^k] \leq \frac{1}{2} k! \sigma^2 b^{k-2}}$

$$E(\exp(\lambda(x-\mu))) \leq 1 + \frac{\lambda^2\sigma^2}{2!} + \frac{\lambda^2\sigma^2}{2} \sum_{k=3}^{\infty} (|\lambda| b)^{k-2}$$

Now lets evaluate

$\sum_{k=3}^{\infty} (|\lambda| 6)^{k-2}$ follows GP sum till infinity

Thus $\sum_{k=3}^{\infty} (|\lambda| 6)^{k-2} = \frac{|\lambda| 6}{1 - |\lambda| 6}$ where $|\lambda| 6 < 1$

i.e. $|\lambda| < \frac{1}{6}$

Putting it back

$$E(\exp(\lambda(x-\mu))) \leq 1 + \frac{\lambda^2 \sigma^2}{2} + \frac{\lambda^2 \sigma^2}{2}\left[\frac{|\lambda| 6}{1 - |\lambda| 6}\right]$$

$\boxed{\text{for } |\lambda| < \frac{1}{6}}$

$$E(\exp(\lambda(x-\mu))) \leq 1 + \frac{\lambda^2 \sigma^2}{2}\left[1 + \frac{|\lambda| 6}{1 - |\lambda| 6}\right]$$

$$\boxed{E(\exp(\lambda(x-\mu))) \leq 1 + \frac{\lambda^2 \sigma^2}{2[1 - |\lambda| 6]}}$$

Now we use the property

$\boxed{1 + t \leq \exp(t)}$ $\forall t \in \mathbb{R}$.

$$E(\exp(\lambda(x-\mu))) \leq 1 + \frac{\lambda^2 \sigma^2}{2[1 - |\lambda| 6]} \leq e^{\frac{\lambda^2 \sigma^2}{2[1 - |\lambda| 6]}}$$

for all $|\lambda| < \frac{1}{6}$.

$$\boxed{E(\exp(\lambda(x-\mu))) \leq e^{\frac{\lambda^2 \sigma^2}{2[1 - |\lambda| 6]}} \quad \forall |\lambda| < \frac{1}{6}}$$

Scanned by CamScanner