Wed_November 25th 2020 Lecture 21 ①

$\Phi_{\vec{X}}(\vec{t}) = E\left[e^{i\vec{t}^T\vec{X}}\right]$

Consider a vector rv x with dimension n. Consider the following operation: $\Phi_{\vec{X}}\left(\begin{bmatrix} t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right) := E\left[e^{i[t\ 0 \dots]\vec{X}}\right] = E\left[e^{itX_1}\right] = \phi_{X_1}(t) \overset{P_1\ P_2}{\Longrightarrow} X_1 \sim f_{X_1}(x)$

$$f_{X_1}(x) = \underbrace{\int_{\mathbb{R}} \cdots \int_{\mathbb{R}}}_{n-1} f_{X_1, X_2, \dots X_n}(x, u_1, \dots, u_{n-1})\, du_1 \dots du_{n-1}$$

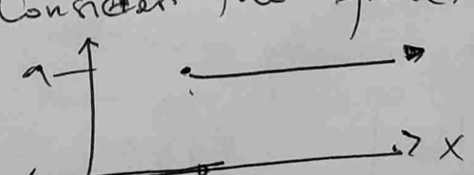
the bottom line is we can use Multivariate chf's to immediately get marginal distribution.

$$\vec{X} \sim N(\vec{\mu}, \Sigma) \Rightarrow \Phi_{\vec{X}}\left(\begin{bmatrix} t \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right) = e^{i[t, 0 \dots 0]\vec{\mu} - \frac{1}{2}[t, 0 \dots 0]\Sigma\begin{bmatrix} t \\ 0 \\ \vdots \\ 0 \end{bmatrix}}$$
$$= e^{it\mu_1 - \frac{t}{2}[t, 0 \dots 0]}$$
$$= e^{it\mu_1 - t^2\sigma_1^2/2} \overset{P_1}{\Longrightarrow} X_1 \sim$$

We now begin the unit on the pure math part of probability with famous inequalities.

Let x be a rv with non-negative support. i.e supp[x] ≥ 0. Let a be a constant >0. Consider the function:

$g(x) := a\mathbb{1}_{x \geq a}$



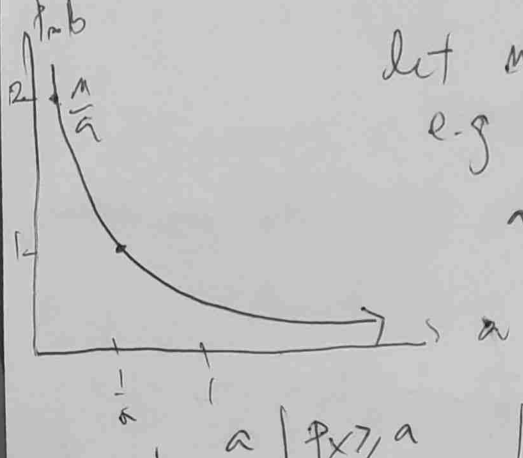
is $a\mathbb{1}_{x \geq a} \overset{?}{\leq} x \quad \forall x$? Consider two cases:

- $x < a \Rightarrow a\mathbb{1}_{x \geq a} = 0 \leq x$ because supp[x] ≥ 0 ✓
- $x \geq a \Rightarrow a\mathbb{1}_{x \geq a} = a \leq x$ because we assume $x \geq a$ ✓

Now let's take the expectation of both sides:

$$E\left(a \mathbb{1}_{X\geq a}\right) \leq E[X] \Rightarrow a\,E\left[\mathbb{1}_{X\geq a}\right] \leq \mu \Rightarrow a\,P(X\geq a) \leq \mu$$

$\Rightarrow P(X\geq a) \leq \frac{\mu}{a}$ this is called Markov's inequality. It's a "tail bound" because it gives you an upper bound on what the probability of the "tail" is. It is a very "crude" bound which means it's seldom so useful and useless if $a < \mu$



let $\mu = 1 = E(X)$

e.g $X \sim Exp(1)$, $P(X\geq a) = 1 - F_X(a) = e^{-a}$

$a = 1 \Rightarrow P(X\geq a) = \frac{1}{e} \simeq 0.37$

a	$P_{X\geq a}$	Markov bound	Chebyshow bound	Chernoff bound
2	0.1353	0.5	1	0.73576
5	0.0067	0.2	.0625	.09158
10	0.00004	0.1	.012	.000123

The Markov inequality has tons of corollaries:

- Let $b = a\mu \Rightarrow P(X\geq b) \leq \frac{\mu}{b} \Rightarrow P(X\geq a\mu) \leq \frac{1}{a}$
- let $h(x)$ be a monotonically increasing function (so it's 1:1).

$$P\left[h(X) \geq h(a)\right] \leq \frac{E\left[h(X)\right]}{h(a)} \Rightarrow P(X\geq a) \leq \frac{E\left[h(X)\right]}{h(a)}$$

- Let $a = \text{Quantile}[X, p] = F_X^{-1}(p)$

$$P(X\geq a) \leq \frac{\mu}{a} \Rightarrow 1 - F_X(a) \leq \frac{\mu}{a} \Rightarrow 1 - F_X\left(F_X^{-1}(p)\right) \leq \frac{\mu}{F_X^{-1}(p)}$$

③ 21

$\Rightarrow 1-p \leq \frac{m}{F_X^{-1}(p)} \Rightarrow F_X^{-1}(p) \leq \frac{m}{1-p}$ e.g. $Med[X] \leq 2m$

. Let X be any r.v $\Rightarrow |X|$ is non negative $\Rightarrow P(|X| \geq a) \leq \frac{E[|X|]}{a}$

. Let X be any rv with finite σ^2 (variance). Let $Y = (X-m)^2 \Rightarrow Y$ is non negative

$$P(Y \geq a^2) \leq \frac{E[Y]}{a^2} \Rightarrow P((X-m)^2 \geq a^2) \leq \frac{E[(X-m)^2]}{a^2}$$

$$\Rightarrow P(|X-m| \geq a) \leq \frac{\sigma^2}{a^2} \quad \text{Chebyshev's Inequality}$$

So let's assume X is non negative and let's get this bound in a more user-friendly form.

$$P(|X-m| \geq a) = P(X-m \geq a \cup -(X-m) \geq a) = P(X-m \geq a) + P(-(X-m) \geq a)$$

$$= P(X-m \geq a) + P(X \leq m-a) \xrightarrow{\text{if } a \geq m}$$ Second term is zero since X is assumed non-negative

$$\Rightarrow P(|X-m| \geq a) = P(X-m \geq a) = P(X \geq a+m) \leq \frac{\sigma^2}{a^2} \Rightarrow P(X \geq b) \leq \frac{\sigma^2}{(b-m)^2}$$

. Let X be any rv and $Y = e^{tX} \Rightarrow Y$ in non neg $\forall t$.

$$\Rightarrow P(Y \geq c) \leq \frac{E(Y)}{c} \Rightarrow P(e^{tX} \geq c) \leq \frac{E[e^{tX}]}{c} \leftarrow \text{mgf}$$

let $c = e^{ta}$, $\Rightarrow P(e^{tX} \geq e^{ta}) \leq e^{-ta} M_X(t)$

$\Rightarrow P(tX \geq ta) \leq e^{-ta} M_X(t)$, $\xRightarrow{\text{if } t>0} P(X \geq a) \leq e^{-ta} M_X(t)$ $\forall t > 0 \xRightarrow{t<0} P(X \leq a) \leq e^{-ta} M_X(t)$

Since this works for all t and we are looking for the best i.e the lowest upper bound, then just optimize over t:

$$\Rightarrow P(X \geq a) \leq \min_{t>0} \{e^{-ta} M_X(t)\} \text{ And } P(X \leq a) \leq \min_{t<0} \{e^{-ta} M_X(t)\}$$

let $X \sim Exp(\lambda) \Rightarrow M_X(t) := E[e^{tX}] = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^\infty e^{(t-\lambda)x} dx$

$$= \lambda \frac{1}{t-\lambda}\left[e^{(t-\lambda)x}\right]_0^\infty = \frac{\lambda}{t-\lambda}\begin{cases}\infty-1 & \text{if } t>\lambda\\ 0-1 & \text{if } t<\lambda\end{cases} = \frac{\lambda}{\lambda-t} \text{ for } t<\lambda$$

otherwise the mgf doesnt exist.

for $X \sim Exp(1)$, the $\boxed{\text{Chernoff bound}}$ is

$$P(X \geq a) \leq \min_{t>0}\left\{e^{-ta}M_X(t)\right\} = \min_{t>0}\left\{e^{-ta}\frac{1}{1-t}\right\} \text{ if } t<1$$

$$= \min_{t\in[0,1]}\left\{\overbrace{\frac{e^{-ta}}{1-t}}^{h(t)}\right\} = \frac{e^{-(1-\frac{1}{a})a}}{1-(1-\frac{1}{a})} =$$

$$h'(t) = \frac{(1-t)(-a)e^{-ta}-(e^{-ta})(-1)}{(1-t)^2} = \frac{a(t-1)e^{-ta}+e^{-ta}}{(1-t)^2} = \frac{e^{-ta}(a(t-1)+1)}{(1-t)^2} = 0$$

$$\Rightarrow at-a+1=0 \Rightarrow t_* = \frac{a-1}{a} = 1-\frac{1}{a} \in (0,1) \text{ if } a>1$$

the reason why the chernoff bound is seldom useful. It requires the mgf. the mgf means you have the f(x)/p(x) and if your know these you may have F(x)