LECTURE 21

$\phi_{\vec{X}}(\vec{t}) := E[e^{i\vec{t}^T\vec{x}}]$

Consider a vector rv X with dimension n. Consider the following operation:

$$\phi_{\vec{X}}\left(\begin{bmatrix} t \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right) := E[e^{i[t\,0\cdots0]\vec{x}}] = E[e^{itX_1}] = \phi_{X_1}(t) \underset{P1,\ P8}{\Rightarrow} X_1 \sim f_{X_1}(x)$$

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

The bottom line is we can use multivariate chf's to immediately get marginal distributions.

$$\vec{X} \sim N(\vec{\mu}, \Sigma) \Rightarrow \phi_{\vec{X}}\left(\begin{bmatrix} t \\ \vdots \\ 0 \end{bmatrix}\right) = e^{i[t\,0\cdots0]\vec{\mu} - \frac{1}{2}[t\,0\cdots0]\Sigma\begin{bmatrix} t \\ \vdots \\ 0 \end{bmatrix}}$$

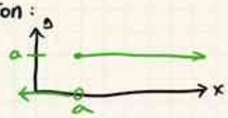
$$= e^{it\mu_1 - \frac{t}{2}[t\,0\cdots0]\begin{bmatrix} \sigma_{11} \\ \vdots \\ \sigma_{n1} \end{bmatrix}}$$

$$= e^{it\mu_1 - t^2\sigma_1^2/2} \overset{P1}{\Rightarrow} X_1 \sim N(\mu_1, \sigma_1^2)$$

We now begin the unit on the "pure math" part of prob. beginning with famous inequalities.

Let X be a rv with non-negative support i.e. $Supp[X] \geq 0$. Let a be a constant > 0.
Consider the function:

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



Is $a\mathbb{1}_{x\geq a} \leq x \quad \forall x$? Consider two cases.

- $x < a \Rightarrow a\mathbb{1}_{x\geq a} = 0 \leq x$ become $Supp[X] \geq 0$
- $x \geq a \Rightarrow a\mathbb{1}_{x\geq a} = a \leq x$ become, we assume, $x \geq a$

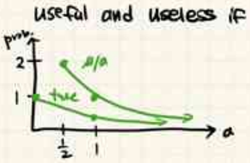
$\Rightarrow a\mathbb{1}_{x\geq a} \leq X$

Now let's take the expectation of both sides:

$E[a\mathbb{1}_{X\geq a}] \leq E[X] \Rightarrow aE[\mathbb{1}_{X\geq a}] \leq \mu \Rightarrow aP(X\geq a) \leq \mu$

$\Rightarrow P(X \geq a) \leq \frac{\mu}{a}$

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. $X \sim Exp(1)$

$P(X \geq a) = 1 - F_X(a) = e^{-a}$

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

a	$P(X\geq a)$	Markov Bound	Chebyshev Bound	Chernoff Bound
2	0.1353	0.5	1	0.73576
5	0.0067	0.2	0.0635	0.09158
10	0.00004	0.1	0.012	0.00123

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)

  $P(h(X) \geq h(a)) \leq \frac{E[h(X)]}{h(a)} \Rightarrow P(X \geq a) \leq \frac{E[h(X)]}{h(a)}$
- let $a = Quantile[X, p] \overset{\text{if } X \text{ continuous}}{=} F_X^{-1}(p)$

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

  $\Rightarrow 1 - p \leq \frac{\mu}{F_X^{-1}(p)} \Rightarrow F_X^{-1}(p) \leq \frac{\mu}{1-p}$ e.g. $Med[X] \leq 2\mu$
- 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 $Y = (X-\mu)^2 \Rightarrow Y$ is non-neg.

$$P(Y \geq a^2) \leq \frac{E[Y]}{a^2} \Rightarrow P((X-\mu)^2 \geq a^2) \leq \frac{\overbrace{E[(X-\mu)^2]}^{\text{def. of variance}}}{a^2}$$

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

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

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

$= P(X-\mu \geq a) + P(-(X-\mu) \geq a) = P(X-\mu \geq a) + P(X \leq \mu - a)$

$\underset{\text{if } a \geq \mu}{\Rightarrow}$ second term is zero since X is assued non-negative

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

$\underset{\text{if } b = \mu + a}{\Rightarrow} P(X \geq b) \leq \frac{\sigma^2}{(b-\mu)^2}$

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

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

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

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

$\underset{\text{if } t>0}{\Rightarrow} P(X \geq a) \leq e^{-ta}M_X(t) \quad \forall t>0 \quad \overset{t<0}{\Longrightarrow} P(X \leq a) \leq e^{-ta}M_X(t)$

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

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

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

$= \lambda\int_0^\infty e^{(t-\lambda)x}dx = \lambda\frac{1}{t-\lambda}[e^{(t-\lambda)x}]_0^\infty = \frac{\lambda}{t-\lambda}\begin{cases} \infty - 1 & \text{if } t>\lambda \\ 0 - 1 & t<\lambda \end{cases}$

$= \frac{\lambda}{\lambda - t}$ for $t < \lambda$

Otherwise the mgf doesn't exist.

For $X \sim Exp(1)$, the Chernoff bound is:

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

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

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

$= \frac{e^{-ta}(a(t-1)+1)}{(1-t)^2} = 0$

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

Let me tell you why the Chernoff bound is seldom useful. It requires the MGF. The MGF means you have the $f(x)$ / $p(x)$ and if you know these you may have $F(x)$ which means you can calculate tail prob. explicitly using numerical integration.