Lecture 21

11/25/20

Math 621
Prof. Kapelner

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

Consider a vector rv X w/ dimension n. Consider the following operation: (P8) (P1)

$\phi_{\vec{X}}\left(\begin{bmatrix} t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right) := E[e^{i[t\,0\ldots0]\vec{X}}] = E[e^{itX_1}] = \phi_{X_1}(t) \Rightarrow$

$X_1 \sim f_{X_1}(x)$

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

$$\vec{X} \sim N(\vec{\mu}, \Sigma) \Rightarrow \phi_{\vec{X}}\left(\begin{bmatrix} t \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right) = e^{i[t\,0\ldots0]\vec{\mu} - \frac{1}{2}[t\,0\ldots0]\Sigma\begin{bmatrix} t \\ 0 \\ \vdots \\ 0 \end{bmatrix}}$$

$$= e^{it\mu_1 - \frac{t}{2}[t\,0\ldots0]\begin{bmatrix} \sigma_{11} \\ \sigma_{21} \\ \vdots \\ \sigma_{n1} \end{bmatrix}}$$

$\rightarrow$ 1

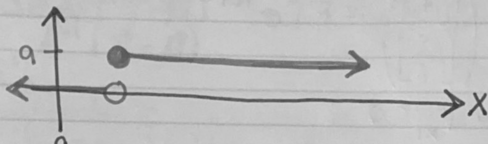
$$= e^{it\mu_1 - t^2\frac{\sigma_1^2}{2}}$$

$$\overset{(P1)}{\Rightarrow} X_1 \sim N(\mu, \sigma_1^2)$$

We now begin the unit on the "pure math" part of prob. beginning w/ famous inequality.

Let X be a rv w/ non-neg. support i.e. $\text{Supp}[X] \geq 0$. Let a be a cts. >0. Consider the function:

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



Is $a\mathbb{1}_{x\geq a} \leq x \ \forall x$?

- if $x < a \rightarrow a\mathbb{1}_{x\geq a} = 0 \leq x$ b/cuz $\text{Supp}[x] \geq 0$. ✓

- if $x \geq a \rightarrow a\mathbb{1}_{x \geq a} = a \leq x$ b/cuz we assume $x \geq a$. ✓

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

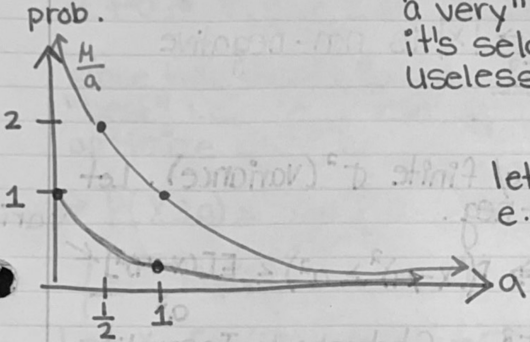
Now let's take the expectation of both sides:

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

this rv has Supp $\{0,1\}$ ⇒ it's a bern.(p)

$\Rightarrow P(X \geq a) \leq \frac{\mu}{a}$ this is called a <u>Markov's</u> <u>inequality</u>

It's a tail bdd. b/cuz it gives you an upper bdd. on what the probability of the "tail" is. It is a very "crude" bdd. which means it's seldom so useful & useless if $a < \mu$.



let $\mu = 1 = E[X]$
e.g. $X \sim Exp(1)$
$P(X \geq a) = 1 - F_X(a) = e^{-a}$

$a = 1 \to P(X \geq a) = \frac{1}{e} \approx .37$

a	$P(X \geq a)$	Markov Bdd.	Chebyshev	Chernoff
2	.1353	.5	1	.73576
5	.0067	.2	.0635	.09158
10	.00004	.1	.012	.00123

The Markov inequalities has ton's of collaries:

- let $b = a\mu \to P(X \geq b) \leq \frac{\mu}{b} \to P(X \geq a\mu) \leq \frac{1}{a}$
- let $h(x)$ be a monotonically increasing fn. (so it's 1:1)
  $P(h(x) \geq h(a)) \leq \frac{E[h(x)]}{h(a)} \to P(X \geq a) \leq \frac{E[h(x)]}{h(a)}$.
  if x cts.
- let $a = \text{Quantile}[X, p] = F_X^{-1}(p)$.

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

$1 - \cancel{F_X}(\cancel{F_X^{-1}}(p)) \leq \frac{\mu}{F_X^{-1}(p)} \Rightarrow 1 - p \leq \frac{\mu}{F_X^{-1}(p)} \Rightarrow F^{-1}(p) \leq \frac{\mu}{1-p}$.

e.g. $\text{Med}[X] \leq 2\mu$

- 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 r.v. w/ finite σ^2 (variance). Let $Y = (X-\mu)^2 \Rightarrow Y$ is non-neg.

variance

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

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

So let's assume X is nonneg. & let's get this bdd. in a more user-friendly form.

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

$= P(X-\mu \geq a) + P(X \leq \mu - a) \rightarrow$ second term is zero since X is assumed non-neg. if $a \geq \mu$

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

let $b = \mu + a$.

$\Rightarrow P(X \geq b) \leq \frac{\sigma^2}{(b-\mu)^2}$

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

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

let $c = e^{ta}$

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

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

if $t>0 \rightarrow P(X \geq a) \leq e^{-ta} M_X(t) \; \forall t>0$

if $t<0 \rightarrow P(X \leq a) \leq e^{-ta} M_X(t)$.

since this works for all t & we are looking for the "best" i.e. the lowest upper bdd., 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 $X \sim Exp(\lambda) \rightarrow M_X(t) := E[e^{tX}] = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx =$

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

$= \frac{\lambda}{\lambda - t}$ for $t<\lambda$ otherwise the mgf does not exist!

For $X \sim Exp(1)$, the chernoff bdd is.....

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

$= \min_{t \in [0,1)} \underbrace{\left\{ \frac{e^{-ta}}{1-t} \right\}}_{= h(t)} = \frac{ae}{e^a}$.

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

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

$\rightarrow at - a + 1 = 0 \rightarrow t_* = \frac{a-1}{a} = 1 - \frac{1}{a} \in (0,1)$

if $a>1$.

Let me tell you why the Chernoff bdd. 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 probabilites explicitly using numerical integration.