Lec 21 Math 621 11/30/17

Markov's Inequality

let X be a non-neg r.v. with finite expectation μ. Consider $a > 0$.
Consider the following inequality.

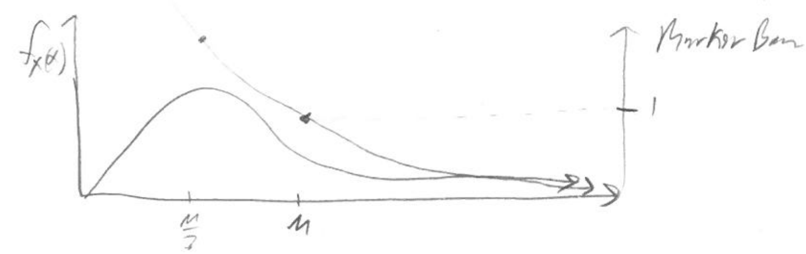
$a \mathbb{1}_{X \geq a} \leq X$. Is this true? Yes...

Why? If $X \geq a \Rightarrow a(1) \leq X \Rightarrow a \leq X \Rightarrow X \geq a$ ✓

If $X < a \Rightarrow a(0) \leq X \Rightarrow 0 \leq X$ true by assumption

Now take expectation of both sides

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



Note $P(X \geq \mu) \leq \frac{\mu}{\mu} = 1$. Not so useful.... More useful for "tail prob's."
i.e. when $a \gg \mu$

Tons of Corollaries / Related Inequalities

* let $a' = a\mu \Rightarrow P(X \geq a\mu) \leq \frac{1}{a}$

* let h be a monotonically increasing function

$$h(a) \mathbb{1}_{h(X) \geq h(a)} \leq h(X) \Rightarrow P\left(h(X) \geq h(a)\right) \leq \frac{E(h(X))}{h(a)}$$

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

* $h(x) = |x|^p$ s.t $p > 1$ are monotonically increasing

$$P(X \geq a) \leq \frac{E[X^p]}{a^p}$$

* let $a = \text{Quantile}[X, p] = F_X^{-1}(p)$ the quantile function

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

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

$$\Rightarrow 1 - F_X(a) \leq \frac{\mu}{a}$$

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

$\Rightarrow 1 - p \leq \frac{\mu}{a} \Rightarrow \text{Quantile}[X, p] \leq \frac{\mu}{1-p}$ relates quantile to expectation!

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

* let X be a r.v. with negative support. Consider $|X|$ which only has pos. support.

$P(|X| \geq a) \leq \frac{E[|X|]}{a}$. All other corollaries apply.

* let X be any r.v. with finite mean and variance

let $Y := (X - \mu_X)^2$. Note: Y has non-neg. support

$P(Y \geq a^2) \leq \frac{E[Y]}{a^2}$ by Markov's inequality

$$\Rightarrow P((X - \mu_X)^2 \geq a^2) \leq \frac{E[(X-\mu_X)^2]}{a^2} = \frac{\sigma^2}{a^2}$$

$\Rightarrow P(|X - \mu_X| \geq a) \leq \frac{\sigma^2}{a^2}$ the difference between a realization and the mean is

Chebyshev's inequality! limited by the variance

~~#~~ let X be any r.v.

$Y = e^{tX}$ Note: Y is always positive

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

let $c = e^{ta}$

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

Note $M_X(t) = E[e^{tX}]$ the moment generating function which is similar to the ch. f.

If $t > 0$

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

$\Rightarrow$ If $t < 0$

$$P(X \leq a) \leq e^{-ta} M_X(t)$$

Since this is valid for all t where the mgf exists...

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

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

Chernoff's Inequality!

Consider $X \sim Bin(n, \frac{1}{4}) \Rightarrow \mu = \frac{1}{4}n$, $\sigma^2 = n\frac{1}{4}(1-\frac{1}{4}) = \frac{3}{16}n$

Note: X is non-negative. We are interested in $P(X > \frac{3}{4}n)$

We know that as n gets large $X \simeq N\left(\frac{1}{4}n, \left(\sqrt{\frac{3}{16}n}\right)^2\right)$ since binomial is sum of iid Bernoullis

$$\Rightarrow P(X > \tfrac{3}{4}n) = P\left(\frac{X - \frac{1}{4}n}{\sqrt{\frac{3}{16}n}} > \frac{\frac{3}{4}n - \frac{1}{4}n}{\sqrt{\frac{3}{16}n}}\right) = P\left(Z > \frac{2}{\sqrt{3}}\sqrt{n}\right) \longrightarrow 0$$ very very quickly.. as n gets larger

$$\frac{\frac{2}{4}}{\sqrt{3n}}$$

Markov: $P\left(X \geq \frac{3}{4}n\right) \leq \frac{\frac{1}{4}n}{\frac{3}{4}n} = \frac{1}{3}$ Note... independent of n.

Chebyshev $P\left(X \geq \frac{3}{4}n\right)$

$$= P\left(X - \tfrac{1}{4}n \geq \tfrac{3}{4}n - \tfrac{1}{4}n\right)$$

why?

$$\leq P\left(X - \tfrac{1}{4}n \geq \tfrac{1}{2}n\right) + P\left(\tfrac{1}{4}n - X \geq \tfrac{1}{2}n\right)$$

$$= P\left(\left|X - \tfrac{1}{4}n\right| \geq \tfrac{1}{2}n\right) \leq \frac{\frac{3}{16}n}{\left(\frac{1}{2}n\right)^2} = \frac{\frac{3}{16}n}{\frac{1}{4}n^2} = \frac{3}{4n} \longrightarrow 0$$

Chernoff's

by binom thm

$$X \sim Bin(n,p) \Rightarrow \phi_X(t) = \sum_{i=0}^{n} e^{itx}\binom{n}{x}p^x(1-p)^{n-x} = \sum_{i=0}^{n}\binom{n}{x}(pe^{it})^x(1-p)^{n-x} = \left(1-p+pe^{it}\right)^n$$

$$M_X(t) = \left(1-p+pe^t\right)^n \qquad X \sim Bin\left(n, \tfrac{1}{4}\right) \Rightarrow M_X(t) = \left(\tfrac{3}{4} + \tfrac{1}{4}e^t\right)^n$$

$$P\left(X \geq \tfrac{3}{4}n\right) \leq \min_{t>0} e^{-t\left(\frac{3}{4}n\right)}\left(\tfrac{3}{4}+\tfrac{1}{4}e^t\right)^n$$

$$= \min_{t>0}\left(e^{-\frac{3}{4}t}\right)^n\left(\tfrac{3}{4}+\tfrac{1}{4}e^t\right)^n$$

$$= \min_{t>0}\left(\tfrac{3}{4}e^{-\frac{3}{4}t} + \tfrac{1}{4}e^{\frac{1}{4}t}\right)^n = \left(\tfrac{3}{4}e^{-\frac{3}{4}\ln(9)} + \tfrac{1}{4}e^{\frac{1}{4}\ln(9)}\right)^n = \left(\tfrac{3}{4}9^{-\frac{3}{4}} + \tfrac{1}{4}9^{\frac{1}{4}}\right)^n$$

$$= \frac{\sqrt[4]{9}}{4^n}\left(\frac{3}{9^3}+1\right)^n = \sqrt[4]{9}\left(\frac{1.004}{4}\right)^n$$

take derivative, set $= 0$

$$\underbrace{n\left(\tfrac{3}{4}e^{-\frac{3}{4}t} + \tfrac{1}{4}e^{\frac{1}{4}t}\right)^{n-1}}_{\text{always positive}}\left(-\tfrac{9}{16}e^{-\frac{3}{4}t} + \tfrac{1}{16}e^{\frac{1}{4}t}\right) = 0$$

$\rightarrow 0$ exponentially fast

$$\Rightarrow e^{\frac{1}{4}t} = 9e^{-\frac{3}{4}t}$$

$$\Rightarrow \tfrac{1}{4}t = \ln(9) - \tfrac{3}{4}t \Rightarrow t^* = \ln(9)$$

Chernoff Bound is the "best"

Consider any two r.v.'s X, Y, with finite mean and variance,

let $W = (X - cY)^2$ s.t. $c \in \mathbb{R}$. Note that W is non-negative

$\Rightarrow E[W] \geq 0$

$E[(X-cY)^2] \geq 0$ Equality? $X = cY$

$E[X^2 - 2cXY + c^2Y^2] \geq 0$

$E(X^2) - 2c\,E[XY] + c^2 E[Y^2] \geq 0$

Since this is valid for all c,

let $c = \dfrac{E(XY)}{E[Y^2]} \in \mathbb{R}$. Pick this to get a nice relationship

$$E[X^2] - 2\frac{E(XY)}{E(Y^2)}E(XY) + \frac{E(XY)^2}{E(Y^2)^{\cancel{2}}}\cancel{E(Y^2)} \geq 0$$

$$E(X^2)E(Y^2) - \cancel{2}E(XY)^2 + \cancel{E(XY)^2} \geq 0$$

$\Rightarrow E[XY]^2 \leq E[X^2]E(Y^2) \Rightarrow |E[XY]| \leq \sqrt{E(X^2)E(Y^2)}$

And if X, Y positive, $E(XY) \leq \sqrt{E(X^2)E(Y^2)}$

Cauchy-Schwarz Inequality

If $X = cY$, what is $Corr[X,Y] = Corr[cY, Y] = \dfrac{Cov[cY,Y]}{SE[cY]\,SE[Y]} = \dfrac{c\,Cov[Y,Y]}{|c|SE(Y)\,SE(Y)}$

$$= \frac{c}{|c|}\frac{Var(Y)}{Var(Y)} = \begin{cases} 1 & \text{if } c > 0 \\ -1 & \text{if } c < 0 \end{cases}$$