Stat 134 lec 13

last time let $\mu_X = E(X)$

$$Var(X) = E\left((X-\mu_X)^2\right)$$

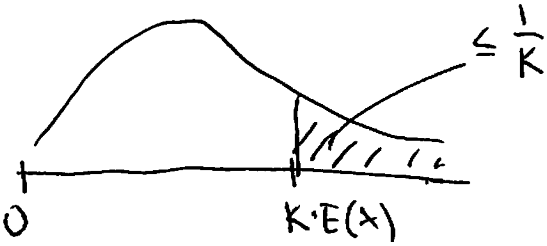
Markov's Ineq

$X \geq 0$, known $E(X)$

$$\boxed{P(X \geq c) \leq \frac{E(X)}{c}}$$

ex

$$P(X \geq \underbrace{5E(X)}_{c}) \leq \frac{\cancel{E(X)}}{5\cancel{E(X)}} = \frac{1}{5}$$

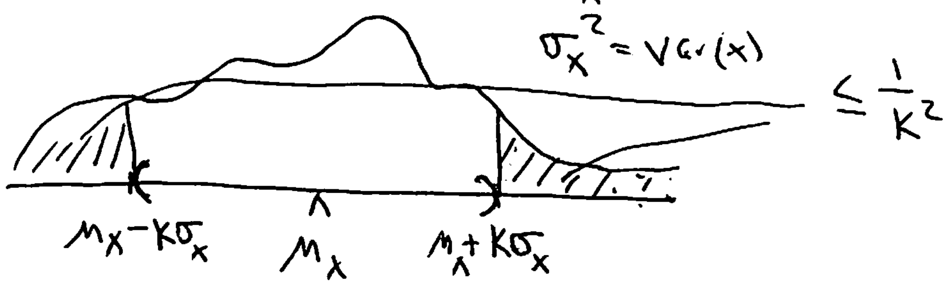


Chebyshev's inequality

$\mu_X = E(X)$

$\sigma_X = SD(X)$ — finite number

$\sigma_X^2 = Var(X)$



$$P(|X-\mu_X| \geq K\sigma_X) = P\big(\underbrace{(X-\mu_X)^2}_{\geq 0} \geq (K\sigma_X)^2\big) \leq \frac{\overbrace{E\left((X-\mu_X)^2\right)}^{\cancel{\sigma_X^2}}}{K^2 \cancel{\sigma_X^2}} = \boxed{\frac{1}{K^2}}$$

ex

Prob to be an outlier (more than 3 SD from avg)

$$\leq \frac{1}{3^2} = \frac{1}{9}.$$

Conversely,

$$P(|X-\mu_X| < 3\sigma_X) \geq 1 - \frac{1}{K^2} = 1 - \frac{1}{9} = \boxed{\frac{8}{9}}$$

If X is normal $P(|X-\mu_X| < 3\sigma_X) = .997$

One sided bound

$P(X \geq \mu_X + k\sigma_X) \leq \frac{1}{k^2}$ since can't assume symmetry

Compare with Markov if $X \geq 0$

$$P(X \geq \mu_X + k\sigma_X) \leq \frac{\mu_X}{\mu_X + k\sigma_X}.$$

I-clicker questions — see next page.

linear change of scale

$$Y = aX + b$$

$$\begin{aligned}
Var(Y) &= E\left((Y - E(Y))^2\right) \\
&= E\left(\left(aX + \cancel{b} - (aE(X) + \cancel{b})\right)^2\right) \\
&= E\left((aX - aE(X))^2\right) \\
&= a^2 E\left((X - E(X))^2\right) = a^2 Var(X)
\end{aligned}$$

$$SD(Y) = \underbrace{\sqrt{a^2}}_{=|a|}\ \underbrace{\sqrt{Var(X)}}_{=SD(X)} = \boxed{|a|\,SD(X)}$$

Another formula for Var(X)

$$\begin{aligned}
Var(X) &= E\left((X - E(X))^2\right) \quad \text{Now FOIL} \\
&= E\left(X^2 - 2XE(X) + (E(X))^2\right) \\
&= E(X^2) - 2E(X)E(X) + E(X)^2 \\
&= \boxed{E(X^2) - (E(X))^2}
\end{aligned}$$

$$2XE(X) = \underbrace{2E(X)}_{c} \cdot X$$

$$E(cX) = cE(X)$$

②

Stat 134

Wednesday February 14 2018

1. A list of incomes has an average of \$70,000 and an SD of \$30,000. Let p be the proportion of incomes over \$100,000. To get an upper bound for p, you should:

   a Assume a normal distribution

   b Use Markov's inequality

   c Use Chebyshev's inequality

   d none of the above

Markov $P(X \geq 100) \leq \frac{70}{100}$

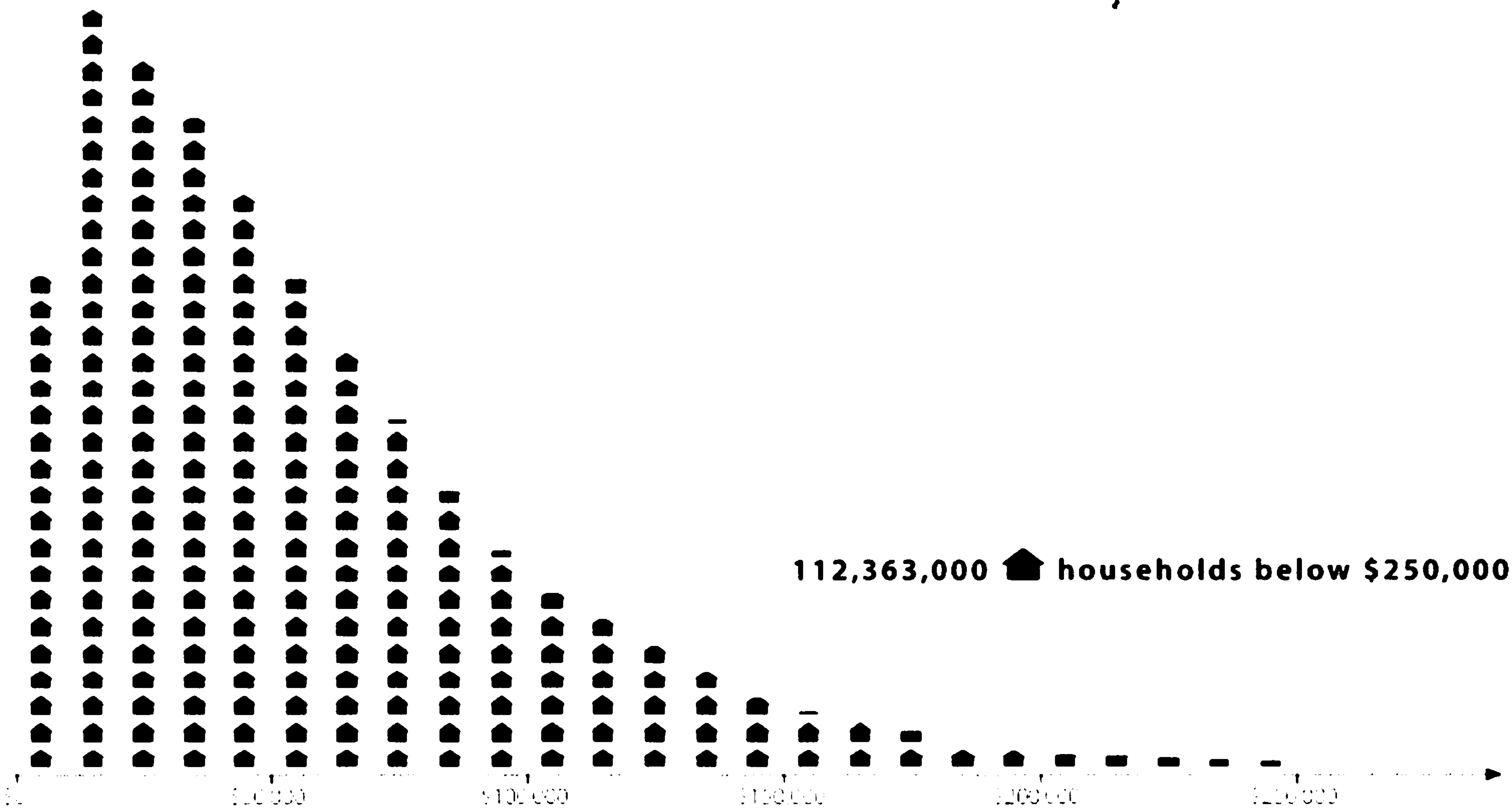
Cheb $P(X > 100) \leq 1 \leftarrow$ not $1/2$

2005 United States

Income Distribution (Bottom 98%)

Each ☗ equals 500,000 households

income isn't normally distributed!



4

2. Suppose there is a large set of data with mean 50 and standard deviation 10. The smallest symmetric interval about the mean that is certain to contain at least 75% of the data points is:

   a (30,70)

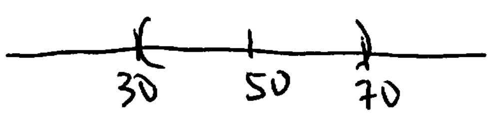
   b (41.3, 58.7)

   c cannot be determined from the information given

   d none of the above

Cheby

$$P(|x-50| < k \cdot 10) \geq 1 - \left(\frac{1}{k^2}\right) \quad \frac{1}{4} \quad \Rightarrow \underline{k=2}$$



Variance of Indicator

I :

0	1
$1-p$	p

$E(I) = p$

$I^2 = I$

$$Var(I) = E(\underset{I}{\underset{\parallel}{I^2}}) - (E(I))^2$$

$$= E(I) - (E(I))^2$$

$$= p - p^2 = \boxed{p(1-p)}$$

Variance Uniform $\{1, 2, \ldots, n\}$

$$E(U) = \frac{n+1}{2}$$

$$E(U^2) = \sum_{k=1}^{n} k^2 \cdot \frac{1}{n} = \frac{(n+1)(2n+1)}{6}$$

$$\frac{n(n+1)(2n+1)}{6}.$$

$$Var(U) = E(U^2) - E(U)^2 = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2$$

— check.

$$\boxed{\frac{n^2-1}{12}}$$

ex

X : roll of a die

$E(X) = 3.5 \quad SD(X) = \sqrt{\frac{36-1}{12}} = \sqrt{\frac{35}{12}}$

Y : unif on $\{17, 18, 19, 20, 21, 22\}$

Find $E(Y)$ $\quad Y = X + 16$

$SD(Y)$. $\quad E(Y) = E(X) + 16 = 3.5 + 16 = 19.5$

$SD(Y) = SD(X) = \sqrt{\frac{35}{12}}$

Several Variables

<u>Recall</u> $E(g(X)) = \sum_{\text{all } x} g(x) P(X=x)$

$$\boxed{E(g(X,Y)) = \sum_{\text{all } x} \sum_{\text{all } y} g(x,y) P(X=x, Y=y)}$$

<u>ex</u> X, Y indep

$$E(XY) = \sum_{\text{all } x} \sum_{\text{all } y} xy P(X=x, Y=y)$$

$$= \sum_{\text{all } x} \sum_{\text{all } y} xy P(X=x) P(Y=y) \quad \text{by indep}$$

$$= \sum_{\text{all } x} \sum_{\text{all } y} \boxed{xP(X=x)}\, y P(Y=y)$$

$$= \sum_{\text{all } x} xP(X=x) \sum_{\text{all } y} yP(Y=y) = E(X)E(Y)$$

so $\boxed{E(XY) = E(X)E(Y) \quad \forall\, X, Y \text{ indep}}$

ex. Take $X=Y$ we get $E(X^2) = (E(X))^2$ — Is this true?