STAT 3600 F 2014 12/02/14

Chebyshev's Inequality:

$K \geq 1$, $\Pr.(|X-\mu| \geq K\sigma) \leq 1/K^2$ for any rv X.

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx = \underbrace{\int_{-\infty}^{\mu-K\sigma} (x-\mu)^2 f(x)\,dx}_{x<\mu-K\sigma} + \int_{\mu-K\sigma}^{\mu+K\sigma} + \underbrace{\int_{\mu+K\sigma}^{\infty} (x-\mu)^2 f(x)\,dx.}_{x>\mu+K\sigma}$$

$$\sigma^2 \geq \int_{-\infty}^{\mu-K\sigma} (x-\mu)^2 f(x)\,dx + \int_{\mu+K\sigma}^{\infty} (x-\mu)^2 f(x)\,dx$$

$$\geq \underbrace{\int_{-\infty}^{\mu-K\sigma} K^2\sigma^2 f(x)\,dx}_{K^2\sigma^2 \Pr.(X<\mu-K\sigma)} + \int_{\mu+K\sigma}^{\infty} K^2\sigma^2 f(x)\,dx = K^2\sigma^2 \Pr(X>\mu+K\sigma)$$

$$\sigma^2 \geq K^2\sigma^2 \underbrace{\left[\Pr.(X-\mu<-K\sigma) + \Pr.(X-\mu>K\sigma)\right]}$$

$$1/K^2 \geq \Pr.(|X-\mu| \geq K\sigma) \qquad \text{QED.}$$

Example. If $E(X) = 25$, $\sigma_X^2 = 3$, then

$\Pr.(|X-25| \geq 1.5\sqrt{3}) \leq 1/1.5^2 = 0.444\bar{4}$

or $\Pr.(\mu - 1.5\sigma \leq X \leq \mu + 1.5\sigma) \geq 0.5555\bar{5}$

Similarly, $\Pr.(|X-25| \geq 3\sigma) \leq 1/3^2 = 1/9$

or: $\Pr.(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \geq 1 - \frac{1}{9} = 8/9$

If $X \sim N(\mu, \sigma^2)$, then $\Pr.(|X-\mu| \leq 3\sigma_X) = 0.9973 >> 8/9$.

If $X \sim N(25, 3.00)$, then the

$$Pr.(|X-\mu| \geq 1.5\sqrt{3}) = Pr.\left(|Z| \geq \frac{1.5\sqrt{3}}{\sqrt{3}}\right) =$$

$$= Pr.(|Z| \geq 1.5) = 2(0.06681) = 0.1336144$$

Compared to 0.4444

If $X \sim U(22, 28)$, then

$$Pr.(|X-25| \geq 1.5\sqrt{3}) = \frac{2}{6}(28-27.5981)$$

$$= 0.1339746..$$

*

Let $\epsilon = k\sigma \longrightarrow Pr.(|X-\mu| \geq \epsilon) \leq 1/(\epsilon/\sigma)^2$

$$\longrightarrow Pr.(|X-\mu| \geq \epsilon) \leq \sigma^2/\epsilon^2.$$

$$Pr.(|\bar{x}-\mu| \geq \epsilon) \leq \sigma_{\bar{x}}^2/\epsilon^2 = \frac{\sigma^2}{n\epsilon^2}$$

$$\lim_{n\to\infty} Pr.(|\bar{x}-\mu| \geq \epsilon) \leq \lim_{n\to\infty} \frac{\sigma^2}{n\epsilon^2} = 0.$$

$\longrightarrow$ This is the law of "Large-Numbers".

similarly, we can show that if $X \sim Bin(n,p)$, then

$$Pr.\left(\left|\frac{X}{n}-p\right| \geq \epsilon\right) \leq \frac{\sigma^2}{\epsilon^2} = \frac{pq/n}{\sigma^2}, \text{ or}$$

$$Pr.\left(\left|\frac{X}{n}-p\right| < \epsilon\right) \geq 1 - \frac{pq}{n\sigma^2} \geq 1 - \frac{1/4}{n\sigma^2} \longrightarrow$$

$$\lim_{n\to\infty}\left(\left|\frac{X}{n}-p\right| < \epsilon\right) \geq \lim_{n\to\infty}\left(1 - \frac{1}{4n\sigma^2}\right) = 1$$

Application of Chebyshev's Inequality

For any rv, X, $Pr.(|X-\mu| \geq K\sigma) \leq 1/K^2$, $K \geq 1$.

Let $K\sigma = \epsilon \rightarrow K = \epsilon/\sigma \rightarrow Pr.(|X-\mu| \geq \epsilon) \leq \dfrac{\sigma_X^2}{\epsilon^2}$,

or: $Pr.(|X-\mu| \leq \epsilon) \geq 1 - \dfrac{\sigma_X^2}{\epsilon^2} \longrightarrow$

$$Pr.(|\bar{X}-\mu| \leq \epsilon) \geq 1 - \frac{\sigma_{\bar{X}}^2}{\epsilon^2} = 1 - \frac{\sigma_X^2}{n\epsilon^2}.$$

As an Example, it is only known that $\sigma_X = 0.40$; then how large of an n do we need such that $\bar{x}$ will be within 0.20 of μ with 98% probability, i.e., determine n such that

$$Pr.(|\bar{X}-\mu| \leq 0.20) \geq 0.98 \rightarrow 1 - \frac{\sigma_X^2}{n\epsilon^2} = 0.98 \rightarrow$$

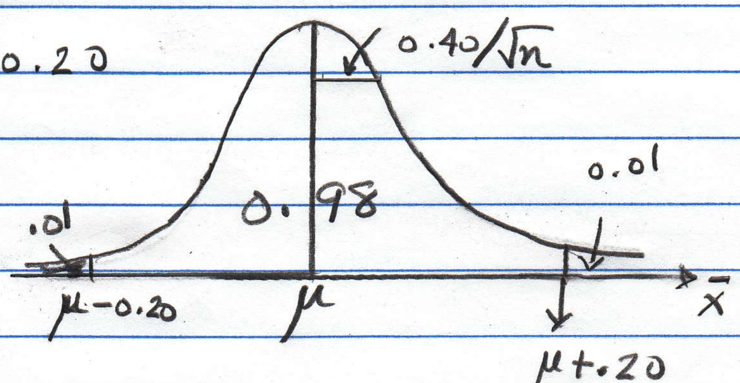
$$1 - \frac{0.16}{n(0.20)^2} = 0.98 \rightarrow n_{min} = 200.$$

Suppose now we have the extra info that $X \sim N(\mu, 0.16)$; then what is the needed n_{min}?

$z_{0.01} \times 0.40/\sqrt{n} \leq 0.20$

$\rightarrow \sqrt{n} \geq 2 z_{0.01}$

$\sqrt{n} \geq 2(2.32635)$



$\sqrt{n} \geq 4.6527 \rightarrow n \geq 21.65 \rightarrow n_{min} = 22.$

Distributions of functions of rv.

Example. Suppose $X \sim U(-2, 3)$ and let $y = x^2$, $0 \le y \le 9$. $\longrightarrow E(X) = 0.50$

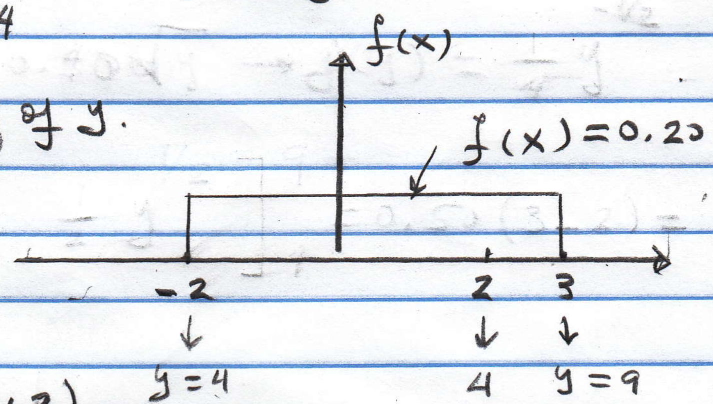
$0 \le y \le 4$

$$G(y) = Pr.(Y \le y) = Pr(X^2 \le y)$$

$$= Pr(-\sqrt{y} \le X \le \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} 0.20\,dx =$$

$$= 0.40\sqrt{y}, \quad 0 \le y \le 4$$

$4 \le y \le 9$; $g(y)$ = The pdf of y.



$$G(y) = Pr(X \le \sqrt{y})$$

$$= \int_{-2}^{\sqrt{y}} 0.20\,dx = 0.20(\sqrt{y} + 2)$$

$$\rightarrow g(y) = 0.10\,y^{-1/2}$$

$$g(y) = \begin{cases} 0.20\,y^{-1/2}, & 0 \le y \le 4 \\ 0.10\,y^{-1/2}, & 4 \le y \le 9 \end{cases}$$

$$E(Y) = \int_0^4 0.20\,y^{1/2}\,dy + \int_4^9 0.10\,y^{1/2}\,dy$$

$$= 0.20\left[\frac{2}{3}y^{3/2}\right]_0^4 + 0.10\left[\frac{2}{3}y^{3/2}\right]_4^9$$

$$= \frac{0.40}{3}(8) + \frac{0.20}{3}(27 - 8) = 7/3$$, or :

$$E(Y) = \int_{R_X} x^2 f(x)\,dx = \int_{-2}^{3} x^2 (0.20)\,dx \qquad (4)$$

$$= 0.20 \left[x^3/3 \right]_{-2}^{3} = 0.20(9 + 8/3) = 7/3, \text{ check!}$$

<u>HW for Thurs. 12/4/2014</u>

Suppose $X \sim U(-2, 3)$ and let $y = x^3$.

(a) obtain $g(y)$ and use it to compute $E(Y)$.

(b) Use $f(x)$ to obtain $E(X^3)$. ANS: $E(X^3) = 3.25$.

*

By def. a χ^2_r is a special case of the Gamma pdf with rate-parameter $\lambda = 1/2$ and $\alpha = n = r/2$, where r is the df of χ^2.

$$\text{gammapdf} = \frac{\lambda}{\Gamma(\alpha)} (\lambda y)^{\alpha - 1} e^{-\lambda y}$$

$$f(\chi^2_r = y) = \frac{1/2}{\Gamma(r/2)} \left(\tfrac{1}{2} y\right)^{\frac{r}{2} - 1} e^{-y/2}, \quad 0 \le y < \infty.$$

$$= \frac{y^{r/2 - 1}}{2^{r/2}\,\Gamma(r/2)} e^{-y/2}, \quad 0 \le y < \infty,$$

where $\Gamma(1/2) = \sqrt{\pi} \rightarrow f(\chi^2_1) = \dfrac{y^{-1/2} e^{-y/2}}{\sqrt{2\pi}}$

How do we generate a χ_1^2 ?

Let $y = z^2$, where $z \sim N(0,1)$; $dy = 2z\,dz$

Then $G_Y(y) = Pr(Y \le y) = Pr(z^2 \le y)$

$$= Pr(-\sqrt{y} \le z \le \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{e^{-z^2/2}}{\sqrt{2\pi}}\,dz$$

$$= \Phi(\sqrt{y}) - \Phi(-\sqrt{y}) = \Phi(\sqrt{y}) - [1 - \Phi(\sqrt{y})]$$

$$= 2\Phi(\sqrt{y}) - 1 \rightarrow g(y) = 2\frac{d}{dy}\Phi(\sqrt{y}) \rightarrow$$

$$g(y) = 2\left[\frac{dz}{dy}\frac{1}{dz}\Phi(z)\right] = 2\left[\frac{1}{2z}\phi(z)\right] =$$

$$= \frac{1}{z}\frac{1}{\sqrt{2\pi}}e^{-z^2/2} = \frac{1}{\sqrt{y}\sqrt{2\pi}}e^{-y/2} = \frac{y^{-1/2}e^{-y/2}}{\sqrt{2\pi}}$$

which is χ^2 with $\nu = df = 1$.

Ind.* Chi-squared rv's are additive; for example,

$\chi_{11}^2 + \chi_{13}^2 \sim \chi_{24}^2$ iff χ_{11}^2 and χ_{13}^2 are independent.

Suppose z_i's are NID(0,1); then $\sum_{i=1}^{n} z_i^2 \sim \chi_n^2$.

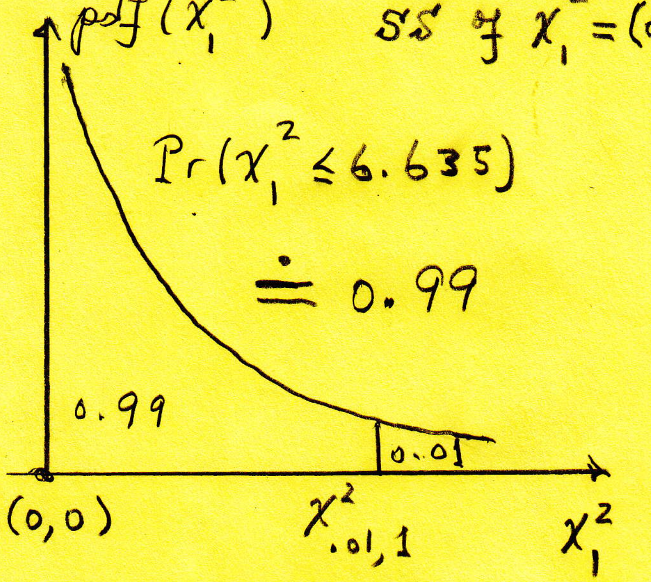
As an example, suppose we draw a random sample of size $n = 20$ from a $N(100, 4)$; then $\sum_{i=1}^{20}\left(\frac{x_i - 100}{2}\right)^2 \sim \chi_{20}^2$ and

$$Pr\left[\sum_{i=1}^{20}\left(\frac{x_i - \mu}{\sigma}\right)^2 \le 28.412\right] = 0.90 \text{ (See Table A-11)}$$

STAT 3600 F 2014 Pictures of χ^2 density

By: S. Maghsoodloo

SS of $\chi_1^2 = (0, \infty)$



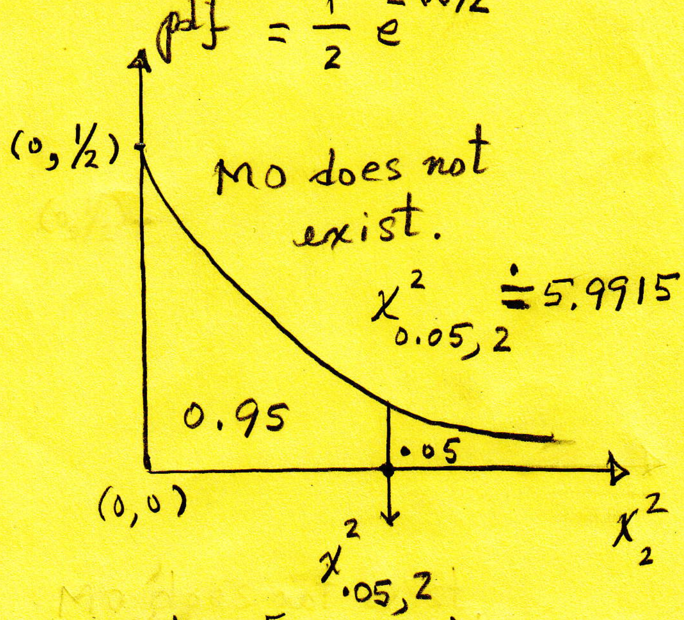
$$C = \frac{1}{2^{\nu/2}\,\Gamma(\nu/2)} \; ; \nu = df$$

At $\nu = 1$, $C = \left(\sqrt{2\pi}\right)^{-1}$

$$f_{\chi_1^2}(w) = C\,w^{-1/2}\,e^{-w/2}$$

$$f'_{\chi_1^2}(w) = -\tfrac{1}{2}\,C\,e^{-w/2}\,w^{-3/2}\,(1+w) \rightarrow \text{the Mode does not exist!}$$

$\chi^2_{0.01,1} = 6.635$ (Matlab) (Minitab); Devore's ANS. on A-11 is inaccurate.



supset = $[0, \infty)$