

Supplemental Notes

EE503 Week 11

Dr. Franzke

HW # 11

① Gubner

6.27 - 6.28

② Leon-Garcia

4.102 - 4.106

5.99 - 5.102

7.23 - 7.29

8.39 - 8.41

Topics

① Central Limit Theorem (CLT)

- $Z_n \xrightarrow{d} Z \sim N(0,1)$ random sample

- $(1-\alpha)$ % c.I. for μ_X : $\bar{X}_n \pm z_{\alpha/2} \cdot \frac{\sigma_X}{\sqrt{n}}$

$$\sigma_{\bar{X}_n} = \text{STD}(\bar{X}_n) = \text{STD}\left(\sum_{k=1}^n X_k\right)$$

Recall for random samples: $X_1, X_2, \dots$ ($\sigma^2 < \infty$)

$$\underline{\text{LLN}}: \bar{X}_n \xrightarrow{\text{mopd}} \mu_x$$

$$\underline{\text{CLT}}: Z_n \xrightarrow{d} Z \sim N(0, 1)$$

$$\text{for } Z_n = \text{STD}(\bar{X}_n) = \frac{\bar{X}_n - \mu_x}{\sigma_x / \sqrt{n}}$$

$$= \text{STD}\left(\sum_{k=1}^n X_k\right)$$

$$= \frac{\sum_{k=1}^n X_k - n \cdot \mu_x}{\sqrt{n} \cdot \sigma_x}$$

$\therefore$ for "large n " ($n \geq 30$):

$$\bar{X}_n \xrightarrow{d} N\left(\mu_x, \frac{\sigma_x^2}{n}\right) \quad \text{and} \quad \sum_{k=1}^n X_k \sim N(n \cdot \mu_x, n \cdot \sigma_x^2)$$

$$\therefore \sqrt{n}(\bar{X}_n - \mu_x) = \frac{\sigma_x}{\sigma_x} \cdot \sqrt{n}(\bar{X}_n - \mu_x)$$

$$= \sigma_x \cdot \left(\frac{\bar{X}_n - \mu_x}{\sigma_x / \sqrt{n}} \right)$$

$$\xrightarrow{d} \sigma_x \cdot Z$$

$$\therefore \sqrt{n}(\bar{X}_n - \mu_x) \xrightarrow{d} N(0, \sigma_x^2)$$

For iid k -vectors $X_k \sim K_{xx}$ $\leftarrow$ positive semi-definite

$$\sqrt{n}(\bar{X}_n - \mu_x) \xrightarrow{d} N\left(0, K_{xx}\right) \quad \text{and} \quad \bar{X}_n \xrightarrow{d} N\left(\mu_x, \frac{1}{n} K_{xx}\right)$$

$k \times 1$ $k \times 1$ $k \times 1$ $k \times k$ $k \times 1$ $k \times 1$ $k \times k$

How good is the CLT approximation?

The Berry-Essen Theorem says that the approximation

error falls off with $\frac{1}{\sqrt{n}}$ if third moment exists.

Thm: (Berry-Essen) $\exists c > 0 \forall x \forall n$.
 $\uparrow$ universal constant

$$\left| F_n(x) - \Phi(x) \right| \leq \frac{c \cdot E[|X - \mu_x|^3]}{\sqrt{n} \cdot \sigma_x^3} \propto \frac{1}{\sqrt{n}}.$$

if iid $X_1, X_2, \dots$ and $E[|X - \mu_x|^3] < \infty$

$$\text{and } F_n(x) = P\left(\frac{\bar{X}_n - \mu_x}{\sigma_x / \sqrt{n}} \leq x\right)$$

and ϕ is CDF of $Z \sim N(0, 1)$

Today: $c < 0.71$ lower bound: $\frac{3 + \sqrt{10}}{6\sqrt{2\pi}} \approx 0.4097$.

CLT Binomial Approximation

For large n : $b(n, k, p) \stackrel{d}{\approx} N\left(\frac{\mu_x}{np}, \frac{\sigma_x^2}{n \cdot pq}\right)$ if $q = 1 - p$.

Rule of thumb: (1) $np \geq 5$
and (2) $nq \geq 5$

Note: $P(3 \leq X \leq 6) \approx P(2.5 \leq X \leq 5.5)$ continuity correction.

$$P(3 \leq X \leq 6) \approx P(2.5 \leq X \leq 6.5)$$

Ex: $X \sim b(10, \frac{1}{2})$

$\therefore n = 10 \quad p = \frac{1}{2}$

$\therefore np = n(1-p) = 10 \cdot \frac{1}{2} = 5.$

$\therefore$ OK to use CLT.

(but NOT Poisson Law)

$\therefore P(3 \leq X \leq 6) \approx P(2.5 \leq X \leq 5.5)$

continuity
correction

standardise

$= P\left(\frac{2.5 - np}{\sqrt{npq}} \leq \frac{X - np}{\sqrt{npq}} \leq \frac{5.5 - np}{\sqrt{npq}}\right)$

$= P\left(\frac{2.5 - 5}{\sqrt{10/4}} \leq \frac{X - 5}{\sqrt{10/4}} \leq \frac{5.5 - 5}{\sqrt{10/4}}\right)$

CLT

$\approx P(-1.581 \leq Z \leq 0.3162)$

for $Z \sim N(0, 1)$

$= \Phi(0.316) - \Phi(-1.581)$

$\approx 0.5684.$

Exact: $P(3 \leq X \leq 6) = 0.5683$ if $X \sim b(10, \frac{1}{2})$

$\therefore |0.5684 - 0.5683| = 0.0001$

Confidence Intervals

2-sided for parameter Θ :

$$\hat{\Theta}_n \pm \text{M.O.E.}$$

M.O.E. = Margin of Error

usually involves $\sigma_{\hat{\Theta}_n} = \sqrt{V[\hat{\Theta}_n]}$
"standard error"

or its estimate $S_{\hat{\Theta}_n}$

$\hat{\Theta}_n$ usually involves $\bar{X}_n$, S_n^2 , or S_{xy}

(random sets $X: \Omega \rightarrow \mathcal{B}$).

Ex: $(1-\alpha)\%$ C.I. for μ_x if σ_x^2 known and

random sample and large $n \geq 30$

$$\text{C.I.} = \bar{X}_n \pm \text{M.O.E.}$$

$$= \bar{X}_n \pm z_{\alpha/2} \left(\frac{\sigma_x}{\sqrt{n}} \right) \leftarrow \sqrt{V[\bar{X}_n]} \text{ standard error}$$

unknown σ_x^2 :

$$\text{C.I.} = \bar{X}_n \pm t_{\alpha/2}(n-1) \frac{S_n}{\sqrt{n}} \leftarrow S_n \text{ estimates } \sigma_x$$

$\uparrow$ wider C.I.

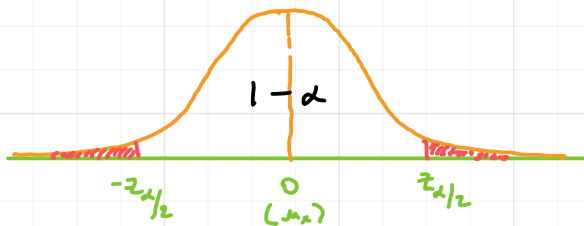
Thrm: $(1-\alpha)\%$ C.I. for μ_x if σ_x^2 known and large n :

$$\bar{X}_n \pm z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}}$$

memorize

(r.s.) $\xrightarrow{\text{IID}} \sigma_x^2 < \infty$

Prf:



$$Z_n \stackrel{\Delta}{=} \frac{X_n - \mu_x}{\sigma_x / \sqrt{n}} \quad \text{since r.s.}$$

$$\therefore \text{large } n: Z_n \stackrel{\text{CLT}}{\approx} Z \sim N(0, 1)$$

$$\therefore \bar{X}_n \stackrel{\text{d}}{\approx} N\left(\mu_x, \frac{\sigma_x^2}{n}\right)$$

$$\therefore 1-\alpha \stackrel{\text{F.T.C.}}{=} P\left(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}\right)$$

$$\stackrel{\text{CLT}}{\approx} P\left(-z_{\alpha/2} \leq \frac{\bar{X}_n - \mu_x}{\sigma_x / \sqrt{n}} \leq z_{\alpha/2}\right) \quad \text{since } n \text{ "large"}$$

$$= P\left(\left\{\omega \in \Omega: -z_{\alpha/2} \leq \frac{\bar{X}_n(\omega) - \mu_x}{\sigma_x / \sqrt{n}} \leq z_{\alpha/2}\right\}\right)$$

$$= P\left(-z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}} \leq \bar{X}_n - \mu_x \leq z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}}\right)$$

$$= P\left(-\bar{X}_n - z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}} \leq -\mu_x \leq -\bar{X}_n + z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}}\right)$$

$$= P\left(\bar{X}_n + z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}} \geq \mu_x \geq \bar{X}_n - z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}}\right)$$

$$= P\left(\bar{X}_n - z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}} \leq \mu_x \leq \bar{X}_n + z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}}\right)$$

$$\therefore \bar{X}_n \pm z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}} \text{ is the } (1-\alpha)\% \text{ CI for } \mu_x \quad \text{QED.}$$

Picking Sample Size n

Want: interval no wider than $\bar{X}_n \pm \epsilon$ in above case

$$\epsilon = \text{Maximal error tolerance} = \text{M.O.E.} = z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}}$$

$$\therefore \epsilon^2 = z_{\alpha/2}^2 \cdot \frac{\sigma_x^2}{n}$$

$$\therefore n = \frac{z_{\alpha/2}^2 \sigma_x^2}{\epsilon^2} \quad \text{memorize}$$

Proportions: Estimate with CLT:

$$X \sim b(n, p) \quad \leftarrow \text{success/failure structure}$$

$$\text{proportion} \quad \hat{p} = \frac{X}{n} \quad \begin{array}{l} \leftarrow \# \text{ success} \\ \leftarrow \# \text{ trials} \end{array}$$

(MLE of p since $X \sim b(n, p)$)

$$\textcircled{1} \quad E[\hat{p}] = \frac{E[X]}{n} = \frac{n \cdot p}{n} = p \quad \forall n \quad \therefore \text{Unbiased}$$

$$\textcircled{2} \quad V[\hat{p}] = \frac{V[X]}{n^2} = \frac{n \cdot p \cdot q}{n^2} = \frac{p \cdot q}{n} \quad \downarrow 0 \text{ as } n \uparrow \infty$$

$$\therefore \hat{p}_n \xrightarrow{P} p \text{ by m.p.d.} \quad \therefore \hat{p}_n \text{ is consistent for } p.$$

$$\therefore \text{For large } n: \text{CLT} \quad \hat{p}_n \stackrel{d}{\approx} N\left(p, \frac{p(1-p)}{n}\right)$$

$\therefore (1-\alpha)\%$ C.I. for p:

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Ex: Polls. " $\pm 3\%$ M.O.E." in the media

implied $\varepsilon = 0.03$ at 95% C.I. ($\alpha = 0.05$)

Note: $p(1-p) \leq \frac{1}{4}$ $\leftarrow$ since $p \in (0,1)$ $\therefore p = \frac{1}{2}$ maximizes $f(p) = p - p^2$

$$\therefore V[\hat{p}] = \frac{p \cdot q}{n} \leq \frac{1}{4n}$$

$\therefore$ Conservative estimate $\cdot \sigma_{\hat{p}} \leq \frac{1}{2\sqrt{n}}$

$$\therefore \varepsilon = \text{M.O.E.} = \frac{z_{\alpha/2}}{2\sqrt{n}}$$

$$\therefore n = \frac{z_{\alpha/2}^2}{4\varepsilon^2}$$

pollster's sample size

$\therefore$ For 95% confidence

Need $n \geq 1068$ samples

For 99% confidence

$n \geq 1849$ samples

"Bootstrap" C.I.

for μ_x (or σ_x or σ_x^2) $\rightarrow$ Bradley-Efron (1979)

Need only iid random samples $X_1, X_2, \dots, X_n$ - so long as they are "statistically representative"

$\therefore$ "non-parametric" (no pdf-assumptions.)
 $\rightarrow$ computer intensive

Ex: Find C.I. for μ_x given 100 r.s. $X_1, \dots, X_{100}$

- ① Form 1000 "resamples" of size 100 each
- sample w/ replacement
 $\rightarrow$ to create "synthetic data"
 $R_1, R_2, \dots, R_{1000}$

- ② Compute the sample mean for each resample R_k .
- $$\begin{array}{ccccccc} R_1 & , & R_2 & , & \dots & , & R_{1000} \\ \downarrow & & \downarrow & & & & \downarrow \\ \bar{X}_{R_1} & , & \bar{X}_{R_2} & , & \dots & , & \bar{X}_{R_{1000}} \end{array}$$

or: $(s_{R_1}^2, s_{R_2}^2, \dots, s_{R_{1000}}^2)$

- ③ Rank-order the 1000 samples

$$\bar{X}_{(1)} \leq \bar{X}_{(2)} \leq \dots \leq \bar{X}_{(25)} \leq \dots \leq \bar{X}_{(975)} \leq \dots \leq \bar{X}_{(1000)}$$

- ④ Pick percentiles 2.5 and 97.5 (for 95% CI):

$$CI = (\bar{X}_{(25)}, \bar{X}_{(975)}) \rightarrow \begin{array}{l} \cdot \text{consistent if "regularity" holds} \\ \cdot \text{do same for } s^2 \text{ or } s, \text{ etc...} \end{array}$$

Thm: If X and Y independent and $Z = X + Y$:

$$f_z = f_x * f_y$$

"splat" of convolution.

Prf: SIT technique (continuous case)

$$F_z(z) \stackrel{\text{CDF}}{=} P[Z \leq z] \stackrel{\text{Hyp}}{=} P[X + Y \leq z]$$

$$\stackrel{\text{total prob}}{=} \int_{x=-\infty}^{x=\infty} P[X + Y \leq z \mid X=x] \cdot f_x(x) dx$$

$$\stackrel{\text{sub}}{=} \int_{x=-\infty}^{x=\infty} P[x + Y \leq z \mid X=x] \cdot f_x(x) dx.$$

$$\stackrel{\text{ind}}{=} \int_{x=-\infty}^{x=\infty} P[x + Y \leq z] f_x(x) dx$$

$$= \int_{x=-\infty}^{x=\infty} P[Y \leq z-x] \cdot f_x(x) dx.$$

$$\therefore f_z(z) = \frac{d}{dz} F_z(z)$$

$$= \frac{d}{dz} \int_{x=-\infty}^{x=\infty} F_Y(z-x) \cdot f_x(x) dx.$$

$$= \int_{x=-\infty}^{x=\infty} \frac{d}{dz} F_Y(z-x) f_x(x) dx \quad (\text{by Leibniz's Rule})$$

$$= \int_{x=-\infty}^{x=\infty} f_Y(z-x) \cdot f_x(x) dx$$

$$= (f_Y * f_x)(z)$$

QED

Defn: Moment Generating Function (MBF)

$$M_X(s) = E_X[e^{sX}] \quad \text{for real } s \text{ where } E_X \text{ exists:}$$

(uniform convergence in discrete case) $-c < s < c$

$s = i\omega$ ($i = \sqrt{-1}$) gives:

Defn: Characteristic function (Lvy, 1920's)

$$\phi_X(\omega) = E_X[e^{i\omega X}]$$

Note $\phi_X(\omega)$ always exists: $|\phi_X(\omega)| \leq E_X[|e^{i\omega X}|] = E[1] = 1$
 $\forall \omega \in \mathbb{R}$

Thm: $E[X^k] = \frac{d^k}{ds^k} M_X(s) \Big|_{s=0}$ "moment generating"

$$- i^k \cdot E_X[X^k] = \frac{d^k}{d\omega^k} \phi_X(\omega) \Big|_{\omega=0}$$

$$\therefore \mu_X = M'_X(0) \quad \sigma_X^2 = M''_X(0) - (M'_X(0))^2$$

Prf: $k=1$ (continuous case when M_X exists)

$$\begin{aligned} \frac{d}{ds} M_X(s) \Big|_{s=0} &= \frac{d}{ds} \int_{x=-\infty}^{x=\infty} e^{sx} f_X(x) dx \Big|_{s=0} \\ &= \int_{x=-\infty}^{x=\infty} \frac{d}{ds} e^{sx} \Big|_{s=0} f_X(x) dx \\ &= \int_{x=-\infty}^{x=\infty} x e^{sx} \Big|_{s=0} f_X(x) dx = \int_{x=-\infty}^{x=\infty} x \cdot e^0 f_X(x) dx \\ &= \int_{x=-\infty}^{x=\infty} x \cdot f_X(x) dx = E_X[X] \end{aligned}$$

Repeat k times for $E[X^k]$.

Put $s = i\omega$ for ϕ_X

QED

Ex: $X \sim b(n, p)$ and $Y \sim P(\lambda)$

$$\begin{aligned}\therefore \mathcal{M}_X(s) &= E_X[e^{sX}] = \sum_{k=0}^n e^{sk} \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=0}^n \binom{n}{k} (pe^s)^k (1-p)^{n-k} \\ &\stackrel{\text{BT}}{=} \boxed{(1-p + pe^s)^n}\end{aligned}$$

$$\begin{aligned}\mathcal{M}_Y(s) &= E_Y[e^{sY}] = \sum_{k=0}^{\infty} e^{sk} \frac{e^{-\lambda} \lambda^k}{k!} \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^s)^k}{k!} \\ &= \boxed{e^{\lambda(e^s - 1)}}$$

★ Thm: (Levy's Continuity Theorem) (use to prove CLT)

$$X_n \xrightarrow{d} X \quad \text{at points of continuity}$$

$$\text{iff } \phi_{X_n} \xrightarrow{c} \phi_X \quad (\text{and } \phi_X \text{ is continuous})$$

Ex: Poisson Law $b \xrightarrow{d} p$ if $\lambda = n \cdot p$

$$\lim_{n \rightarrow \infty} \mathcal{M}_X(s) = \lim_{n \rightarrow \infty} (1-p + pe^s)^n \quad \text{if } X \sim b(n, p)$$

$$\stackrel{\lambda = np}{=} \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n} + \frac{\lambda e^s}{n}\right)^n = \lim_{n \rightarrow \infty} \left(1 + \frac{\lambda(e^s - 1)}{n}\right)^n$$

$$= e^{\lambda(e^s - 1)} = \mathcal{M}_Y(s) \quad \text{if } Y \sim P(\lambda)$$

$\therefore b \xrightarrow{d} p$ by Levy's Theorem



Thrm: If $X \sim N(\mu_x, \sigma_x^2)$

$$1-D \left\{ \begin{aligned} m_x(s) &= e^{s\mu + s^2\sigma^2/2} \\ \phi_x(\omega) &= e^{i\omega\mu - \omega^2\sigma^2/2} \end{aligned} \right.$$

memorize

II+ $X \sim N(\underbrace{\mu_x}_{n \times 1}, \underbrace{K_{xx}}_{n \times n})$ positive semi-definite

$$n-D \left\{ \begin{aligned} m_x(s) &= \exp\left(s^T \mu_x + \frac{s^T K_{xx} s}{2}\right) \\ \phi_x(\omega) &= \exp\left(i\omega^T \mu_x - \frac{\omega^T K_{xx} \omega}{2}\right) \end{aligned} \right.$$

Note: K_{xx}^{-1} need not exist (unlike pdf f_x)
eigenvalues $\lambda = 0$ OK.

Prf: (1-D case)

$$m_x(s) = E_x[e^{sX}] = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{sx} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \quad X \sim N(\mu, \sigma^2)$$

$$\begin{aligned} & \text{put } z = \frac{x-\mu}{\sigma} \quad \therefore x = \sigma z + \mu \quad \begin{matrix} dx = \sigma dz \\ (x \rightarrow \pm\infty \rightarrow z \rightarrow \pm\infty) \end{matrix} \\ & = \frac{1}{\sqrt{2\pi}\cancel{\sigma}} \int_{-\infty}^{\infty} e^{s(\sigma z + \mu)} e^{-z^2/2} \cancel{\sigma} dz \\ & = \frac{e^{s\mu}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2 - 2zs\sigma)} dz \\ & \stackrel{\text{complete the square}}{=} \frac{e^{s\mu}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}[(z-s\sigma)^2 - s^2\sigma^2]} dz. \end{aligned}$$

$$\text{since } (z-s\sigma)^2 - s^2\sigma^2 = z^2 + s^2\sigma^2 - 2zs\sigma - s^2\sigma^2 = z^2 - 2zs\sigma$$

$$= e^{s\mu + \frac{s^2\sigma^2}{2}} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z-s\sigma)^2} dz \right)$$

$$\begin{aligned}
 & \text{put } u = z - s\tau \quad \therefore du = dz, \quad z = \pm \infty \longrightarrow u = \pm \infty \\
 & \stackrel{u=z-s\tau}{=} e^{s\mu + s^2\tau^2/2} \left(\underbrace{\frac{1}{\sqrt{2\pi}} \int_{u=-\infty}^{u=\infty} e^{-u^2/2} du}_{=1 \text{ since pdf}} \right) \\
 & = e^{s\mu + s^2\tau^2/2}
 \end{aligned}$$

QED.

Proposition: If X and Y independent and $Z = X + Y$

$$\begin{cases} m_z = m_x m_y \\ \phi_z = \phi_x \phi_y \end{cases}$$

$$n\text{-d: } m_{\sum_{k=1}^n X_k} = \prod_{k=1}^n m_{X_k}$$

$$\begin{aligned}
 \text{Prf: } m_z(s) &= E_{x,y} [e^{sX+sY}] = E_{x,y} [e^{sX} e^{sY}] \\
 &\stackrel{\text{ind}}{=} E_{x,y} [e^{sX}] \cdot E_{x,y} [e^{sY}] = E_x [e^{sX}] \cdot E_y [e^{sY}] \\
 &= m_x(s) \cdot m_y(s)
 \end{aligned}$$

QED.

$$\text{Ex: Cauchy } X \sim C(m, d): f_x(x) = \frac{1}{\pi d (1 + (\frac{x-m}{d})^2)}$$

m_x does not exist, But ϕ_x does

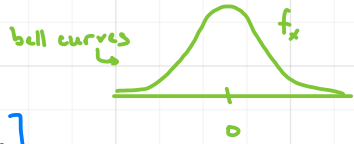
$$\phi_x(\omega) = e^{i\omega m - d|\omega|}$$

from Cauchy's Residue Theorem:
 $Z \sim C(0, 1): \phi_z(\omega) = e^{-|\omega|}$

General Symmetric Alpha Stable (S α S)

(Paul Levy, 1923)

$$\phi_x(w) = e^{-d|w|^\alpha}$$



for tail-thickness $\alpha \in (0, 2]$

No closed form pdf except

$\alpha = \frac{1}{2}$ (Levy) asymmetric
 $\alpha = 1$ (Cauchy)
 $\alpha = 2$ (Gaussian)

$\alpha \downarrow \longleftrightarrow$ bell-curve tail thickness $\uparrow$

$E_x[|X|^f]$ exists if $f < \alpha$
(lower-order Fractional moments)

Higher-order moments do not exist if $\alpha < 2$
 $E[|X|^f] = \infty$ if $f \geq \alpha$

General Stable Distributions

$$\phi_x(w) = \begin{cases} \exp[-\gamma^\alpha |w| [1 + \beta \tan \frac{\alpha\pi}{2} \cdot \text{sign}(w) (|w|^{1-\alpha} - 1)] + i m w] & \text{if } \alpha \neq 1 \\ \exp[-\gamma |w| [1 + \beta \frac{2}{\pi} (\text{sign}(w) (\log |w|))] + i m w] & \text{if } \alpha = 1 \end{cases}$$

- tail-thickness $0 < \alpha \leq 2$
- dispersion $\gamma > 0$
- location $m \in \mathbb{R}$
- skew $-1 \leq \beta \leq 1$ ($\beta = 0$ symmetric pdf)

★ Inversion (Transform) Theorem

If CDF F_x is absolutely continuous (so pdf f_x exists):

$$f_x(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_x(w) \cdot e^{-iwx} dw$$

Ex: $X \sim C(m, d)$

$$\begin{aligned} \therefore f_x(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iwm - d|w|} e^{-iwx} dw \\ &= \frac{1}{2\pi} \left[\int_0^{\infty} e^{iwm - dw} e^{-iwx} dw + \int_{-\infty}^0 e^{iwm + dw} e^{-iwx} dw \right] \end{aligned}$$

since $|w| = \begin{cases} w & \text{if } w \geq 0 \\ -w & \text{if } w < 0 \end{cases}$

$$= \frac{1}{2\pi} \left[\int_0^{\infty} e^{-w(d+ix-im)} dw + \int_{-\infty}^0 e^{w(d-ix+im)} dw \right]$$

$u = -w(d+ix-im)$
 $\therefore du = -dw(d+ix-im)$
 $\begin{cases} w=0 \rightarrow u=0 \\ w=\infty \rightarrow u=-\infty \end{cases}$

$v = w(d-ix+im)$
 $dv = dw(d-ix+im)$
 $\begin{cases} w=0 \rightarrow v=0 \\ w=-\infty \rightarrow v=-\infty \end{cases}$

$$= \frac{1}{2\pi} \left[-\frac{1}{d+ix-im} \int_{u=0}^{-\infty} e^u du + \frac{1}{d-ix+im} \int_{v=-\infty}^0 e^v dv \right]$$

$$= \frac{1}{2\pi} \left[\left. \frac{e^u}{d+ix-im} \right|_{-\infty}^0 + \left. \frac{e^v}{d-ix+im} \right|_{-\infty}^0 \right]$$

$$= \frac{1}{2\pi} \left[\frac{1}{d+ix-im} + \frac{1}{d-ix+im} \right]$$

$$= \frac{1}{2\pi} \left[\frac{2d}{d^2 + x^2 + m^2 - 2mx} \right] \quad \text{after rationalizing}$$

$$= \frac{1}{\pi d} \cdot \frac{1}{\left(1 + \left(\frac{x-m}{d}\right)^2\right)}$$

$$= f_x(x) \quad \longleftrightarrow \quad X \sim C(m, d)$$

Thrm: If independent $X_1, \dots, X_n$ and $X_k \sim N(\mu_k, \sigma_k^2)$ then

$$\sum_{k=1}^n a_k X_k \sim N\left(\sum_{k=1}^n a_k \mu_k, \sum_{k=1}^n a_k^2 \sigma_k^2\right)$$

Prf: $\phi_{\sum_{k=1}^n a_k X_k}(\omega) = E\left[e^{i\omega \sum_{k=1}^n a_k X_k}\right]$

$$= E\left[\prod_{k=1}^n e^{i\omega a_k X_k}\right]$$

ind

$$= \prod_{k=1}^n E\left[e^{i(\omega a_k) X_k}\right]$$

$$= \prod_{k=1}^n \phi_{X_k}(\omega a_k)$$

$$= \prod_{k=1}^n e^{i\omega a_k \mu_k - \frac{\omega^2}{2} a_k^2 \sigma_k^2}$$

$$= e^{i\omega \left(\sum_{k=1}^n a_k \mu_k\right) - \frac{\omega^2}{2} \left(\sum_{k=1}^n a_k^2 \sigma_k^2\right)}$$

$$\longleftrightarrow \sum_{k=1}^n a_k X_k \sim N\left(\sum_{k=1}^n a_k \mu_k, \sum_{k=1}^n a_k^2 \sigma_k^2\right)$$

QED.

Q: How does $\bar{X}_n$ differ if (iid) $X_k \sim N(\mu, \sigma^2)$ vs. $X_k \sim C(m, d)$

Thrm: If iid $X_1, \dots, X_n \sim N(\mu, \sigma^2)$: $\bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

Prf: $\phi_{\bar{X}_n}(\omega) = E\left[e^{i\omega \bar{X}_n}\right] = E\left[e^{i\omega \frac{1}{n} \sum_{k=1}^n X_k}\right]$ (exact, not CLT approximation)

$$= E\left[e^{i \sum_{k=1}^n \frac{\omega}{n} X_k}\right] = E\left[\prod_{k=1}^n e^{i \frac{\omega}{n} X_k}\right]$$

ind

$$= \prod_{k=1}^n E\left[e^{i \left(\frac{\omega}{n}\right) X_k}\right] = \prod_{k=1}^n \phi_{X_k}\left(\frac{\omega}{n}\right)$$

$X_k \sim N(\mu, \sigma^2)$

$$= \prod_{k=1}^n e^{i \frac{\omega}{n} \mu - \frac{\omega^2}{n^2} \frac{\sigma^2}{2}}$$

$$\begin{aligned}
 \text{Id.} \quad &= \left(e^{i(\frac{w}{n})\mu - \frac{w^2}{2}(\frac{\sigma^2}{n})} \right)^n \\
 &= e^{i w \mu - \frac{w^2}{2} \frac{\sigma^2}{n}} \longleftrightarrow \bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)
 \end{aligned}$$

QED

★
Thm If iid $X_1, \dots, X_n \sim C(m, d)$: $\bar{X}_n \sim C(m, d)$
 $\therefore$ no benefit to averaging!

Prf: $\phi_{\bar{X}_n}(w) = \prod_{k=1}^n E[e^{i(\frac{w}{n})X_k}]$ (as in previous proof)

$$\begin{aligned}
 &= \prod_{k=1}^n \phi_{X_k}\left(\frac{w}{n}\right) \\
 &\stackrel{X_k \sim C(m, d)}{=} \prod_{k=1}^n e^{i(\frac{w}{n})m - d|\frac{w}{n}|}
 \end{aligned}$$

$$= e^{\sum_{k=1}^n (i(\frac{w}{n})m - d|\frac{w}{n}|)}$$

$$\stackrel{\text{Id.}}{=} e^{(i \cdot \cancel{n} \cdot \frac{w}{\cancel{n}} \cdot m - \cancel{n} \cdot d \cdot \frac{|w|}{\cancel{n}})}$$

$$= e^{i w m - d|w|} \longleftrightarrow \bar{X}_n \sim C(m, d)$$

QED

$$\text{Var}[\bar{X}_n] = \frac{\sigma_x^2}{n} \text{ requires } \sigma_x^2 < \infty$$

Note: Francis Edgeworth showed in 1883 that $\sigma_x^2 < \infty$ is not sufficient, even when $n=2$.

Central Limit Theorem ★★

Thm: If iid $X_1, X_2, \dots$ and $\sigma^2 < \infty$ ($\therefore |\mu_k| < \infty$) then

$$Z_n \xrightarrow{d} Z \sim N(0, 1)$$

$$\begin{aligned} \text{where } Z_n &= \text{STD}(\bar{X}_n) = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \\ &= \text{STD}\left(\sum_{k=1}^n X_k\right) = \frac{\sum_{k=1}^n X_k - n \cdot \mu}{\sqrt{n} \cdot \sigma} \end{aligned}$$

Prf: Strategy: show $\mathcal{M}_{Z_n} \xrightarrow{c} \mathcal{M}_Z(s) = e^{s^2/2}$ and then Levy's theorem.

$$\begin{aligned} \mathcal{M}_{Z_n}(s) &= E[e^{sZ_n}] = E\left[e^{s\left(\frac{\sum_{k=1}^n X_k - n \cdot \mu}{\sqrt{n} \cdot \sigma}\right)}\right] \\ &= E\left[e^{\sum_{k=1}^n \frac{s}{\sqrt{n} \cdot \sigma} (X_k - \mu)}\right] \\ &= E\left[\prod_{k=1}^n e^{\frac{s}{\sqrt{n} \cdot \sigma} (X_k - \mu)}\right] \stackrel{\text{ind}}{=} \prod_{k=1}^n E\left[e^{\frac{s}{\sqrt{n} \cdot \sigma} (X_k - \mu)}\right] \\ &= E^n\left[e^{\frac{s(X-\mu)}{\sqrt{n} \cdot \sigma}}\right] \\ &= \left(E\left[1 + \frac{s(X-\mu)}{\sqrt{n} \cdot \sigma} + \frac{s^2(X-\mu)^2}{2n\sigma^2} + \text{H.O.T.}\right]\right)^n \\ &= \left(1 + 0 + \frac{s^2 \sigma^2}{2n\sigma^2} + \text{H.O.T.}\right)^n \\ &= \left(1 + \frac{s^2/2}{n} + \text{H.O.T.}\right)^n \xrightarrow{c} e^{s^2/2} \\ &= \mathcal{M}_Z(s) \longleftrightarrow Z \sim N(0, 1) \end{aligned}$$

since $\lim_{n \rightarrow \infty} \left(1 + \frac{c/n}{n}\right)^n = e^c$ if $\lim_{n \rightarrow \infty} c/n = c$

$\therefore Z_n \xrightarrow{d} Z$ by Levy's continuity theorem, since $e^{s^2/2}$ is continuous
QED

∴ "Delta Method" ^(differentiable) for continuous g (not involving n)

$$\text{If } \sqrt{n}(X_n - \mu) \xrightarrow{d} N(0, \sigma^2)$$

$$\text{then } \sqrt{n}(g(X_n) - g(\mu)) \xrightarrow{d} N(0, \sigma^2 [g'(\mu)]^2)$$

follows from mean-value theorem and continuity of Plim.

"Lindeberg condition" yields the CLT when only independence hold
(iid need not)

Thm: (Lindeberg-Feller CLT)

If - $X_1, X_2, \dots$ are independent and $\sigma_k^2 < \infty$

$$- \lim_{n \rightarrow \infty} \bar{\sigma}_n^2 = \bar{\sigma}^2 < \infty \quad \text{for } \bar{\sigma}_n^2 = \frac{1}{n}(\sigma_1^2 + \dots + \sigma_n^2)$$

$$- \lim_{n \rightarrow \infty} \frac{\max_k \sigma_k}{n \cdot \bar{\sigma}_n} = 0 \quad (\text{i.e. no term dominates the average variance})$$

$$\text{Then: } \sqrt{n}(\bar{X}_n - \bar{\mu}_n) \xrightarrow{d} N(0, \bar{\sigma}^2)$$

$$\text{for } \bar{\mu}_n = \frac{1}{n}(\mu_1 + \dots + \mu_n)$$

Thm: (Generalized CLT), CLT holds $\iff$ stable r.v.

iid $X_1, X_2, \dots$ then $\exists a_n > 0$ and $b_n \in \mathbb{R}$ and non-degenerate

$$\text{r.v. } Z: a_n \sum_{k=1}^n X_k - b_n \xrightarrow{d} Z \text{ iff } Z \text{ is } \alpha\text{-stable for some } \alpha \in (0, 2]$$

$$\therefore \text{Classical CLT if } Z \sim N(0, 1) \text{ and } a_n = \frac{1}{\sqrt{n}\sigma} \text{ and } b_n = \sqrt{n} \frac{\mu}{\sigma} \quad (\alpha=2)$$

Thm: (S&S case of CLT-like averaging) If iid

$$X_1, X_2, \dots, X_n \sim \phi_x(\omega) = e^{-c_\alpha |\omega|^\alpha} \text{ and } 0 < \alpha \leq 2 \text{ then}$$

$$\frac{1}{n^{1/\alpha}} \sum_{k=1}^n X_k \sim \phi_x(\omega) = e^{-c_\alpha |\omega|^\alpha}$$

Already saw this for $\alpha=2$: $\bar{X}_n \sim N(\mu, \frac{\sigma^2}{n})$ and

for $\alpha=1$: $\bar{X}_n \sim C(m, d)$

Extreme Value "CLT"

- extreme value: $P(X > 0)$

$$M_n \triangleq \max(X_1, \dots, X_n)$$

dual result for min since
 $\min(X_1, \dots, X_n) = -\max(X_1, \dots, X_n)$

- M_n degenerates as $n \rightarrow \infty$

- need to find a_n and b_n that normalize M_n

- turns out only 3 limiting distribution

Thm: (Extreme Value Theorem, Fisher-Tippett Theorem) If

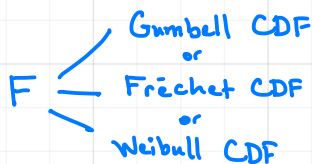
(1) $X_1, X_2, \dots$ are iid

$$(2) \exists a_n > 0 \text{ and } b_n \in \mathbb{R} : \lim_{n \rightarrow \infty} P\left(\frac{M_n - b_n}{a_n} \leq x\right) = F_x(x)$$

and F is not a degenerate CDF for $M_n = \max(X_1, \dots, X_n)$

then:

F must be one of 3 "attractor" CDFs :



Ex: $X_k \sim N(\mu, \sigma^2)$, $X_k \sim \exp(\Theta)$, $X_k \sim \text{lognormal}(\mu, \sigma^2)$

$$\therefore F \sim \text{Gumbel} : F_x(x) = e^{-e^{\left(\frac{x-\mu}{\sigma}\right)}}$$

Ex: $X_k \sim \mathcal{L}(m, d)$, $X \sim \text{SaS}$ for $\alpha \leq 2$, $X_k \sim t(n-1)$ student, $X_k \sim \text{Pareto}$

$$\therefore F \sim \text{Fréchet} : F_x(x) = e^{-\left(\frac{x-\mu}{\sigma}\right)^{-\alpha}} \text{ if } \alpha \in (0, \infty)$$

Ex: $X_k \sim U(a, b)$, $X_k \sim \text{Beta}(\alpha, \beta)$

$$\therefore F \sim \text{Weibull} : F_x(x) = 1 - e^{-(x/\beta)^\alpha}$$

Ex: $X_k \sim \text{Poisson}(\lambda)$

$\therefore$ no attractor CDF (degenerate)

Ex: Pick $a_n = 1 > 0$ and $b_n = \ln n$ and $X_k \sim \exp(1)$ iid

$$\therefore P\left[\frac{M_n - b_n}{a_n} \leq x\right] = P[M_n \leq a_n x + b_n] \stackrel{\text{iid}}{=} P^n(X \leq a_n x + b_n)$$

$$\stackrel{X_k \sim \exp(1)}{=} \left(1 - e^{-a_n x - b_n}\right)^n \stackrel{a_n=1, b_n=\ln n}{=} \left(1 - e^{-x - \ln n}\right)^n \stackrel{\text{since } M_n = \max(X_1, \dots, X_n)}{=} \left(1 - e^{\ln\left(\frac{1}{n}\right)} e^{-x}\right)^n$$

$$= \left(1 - \frac{e^{-x}}{n}\right)^n \xrightarrow{d} e^{-e^{-x}} = F_x(x) \longleftrightarrow$$

Gumbel CDF.

★ Gamma pdf

$$X \sim \mathcal{G}(\alpha, \theta), \quad \therefore f_X(x) = \frac{x^{\alpha-1} e^{-x/\theta}}{\Gamma(\alpha) \theta^\alpha} \quad \text{if } x > 0 \quad (\alpha > 0 \text{ and } \theta > 0)$$

$$\text{for } \Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$$

$$\therefore \Gamma(\alpha+1) = \alpha \cdot \Gamma(\alpha) \quad \text{from integration by parts}$$

Thm: Say $X \sim \mathcal{G}(\alpha, \theta)$

$$\textcircled{1} E[X^k] = \theta^k \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)} \quad \text{if } k > 0$$

moments

$$\therefore E[X] = \alpha \theta \quad V[X] = \alpha \theta^2$$

$$\textcircled{2} \text{ MGF: } \mathcal{M}_X(s) = \frac{1}{(1-s\theta)^\alpha} \quad \text{if } s < \frac{1}{\theta}$$

Additivity. Say $X_1, \dots, X_n$ are independent and

$$X_k \sim \mathcal{G}(\alpha_k, \theta) \quad \text{same } \theta \quad \text{Then } X = \sum_{k=1}^n X_k \sim \mathcal{G}\left(\sum_{k=1}^n \alpha_k, \theta\right).$$

$$\therefore \sum_{k=1}^r X_k \sim \chi^2(r) \quad \text{if iid } X_k \sim \chi^2(1)$$

$$\text{since } X \sim \chi^2(r) \text{ if } X \sim \mathcal{G}\left(\frac{r}{2}, 2\right)$$

Pft:

$$\textcircled{1} E[X^k] = \frac{1}{\Gamma(\alpha) \theta^\alpha} \int_0^\infty x^k x^{\alpha-1} e^{-x/\theta} dx$$

$$= \frac{1}{\Gamma(\alpha) \theta^\alpha} \int_0^\infty x^{(\alpha+k)-1} e^{-x/\theta} dx$$

$$\text{put } u = x/\theta \quad \therefore x = \theta u \quad \text{and } dx = \theta du$$

$$x = 0 \longrightarrow u = 0 \quad \text{since } \theta > 0$$

$$x = \infty \longrightarrow u = \infty$$

$$\begin{aligned}
&= \frac{1}{\Gamma(\alpha)\Theta^\alpha} \int_{u=0}^{u=\infty} \Theta^{(\alpha+k)-1} u^{(\alpha+k)-1} e^{-u} \Theta du. \\
&= \frac{\Theta^{\alpha+k}}{\Gamma(\alpha)\Theta^\alpha} \int_0^\infty \underbrace{u^{(\alpha+k)-1} e^{-u}}_{\Gamma(\alpha+k)} du \\
&= \Theta^\alpha \cdot \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}
\end{aligned}$$

$$\therefore E[X] = \Theta^\alpha \cdot \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} = \Theta^\alpha \cdot \frac{\alpha \cdot \Gamma(\alpha)}{\Gamma(\alpha)} = \alpha \Theta.$$

$$E[X^2] = \Theta^\alpha \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} = \Theta^\alpha \cdot \frac{(\alpha+1)\alpha\Gamma(\alpha)}{\Gamma(\alpha)} = \Theta^\alpha \alpha(\alpha+1)$$

$$= \alpha^2 \Theta^2 + \alpha \Theta^2$$

$$\therefore V[X] = E[X^2] - E^2[X] = \cancel{\alpha^2 \Theta^2} + \alpha \Theta^2 - \cancel{\alpha^2 \Theta^2} = \alpha \Theta^2.$$

QED (1)

$$\begin{aligned}
\textcircled{2} \quad \mathcal{M}_x(s) &= E[e^{sX}] = \frac{1}{\Gamma(\alpha)\Theta^\alpha} \int_0^\infty e^{sx} x^{\alpha-1} e^{-x/\Theta} dx \\
&= \frac{1}{\Gamma(\alpha)\Theta^\alpha} \int_0^\infty x^{\alpha-1} e^{-x(\frac{1}{\Theta}-s)} dx \quad \text{if } s < \frac{1}{\Theta} \\
&= \frac{1}{\Gamma(\alpha)\Theta^\alpha} \int_0^\infty x^{\alpha-1} e^{-x(\frac{1-s\Theta}{\Theta})} dx
\end{aligned}$$

$$\text{put } u = x \frac{(1-s\Theta)}{\Theta} \longleftrightarrow x = \frac{\Theta u}{1-s\Theta} \quad \text{and } dx = \frac{\Theta du}{1-s\Theta}$$

$$= \frac{1}{\Gamma(\alpha)\Theta^\alpha} \int_{u=0}^{u=\infty} \frac{\Theta^{\alpha-1} u^{\alpha-1}}{(1-s\Theta)^{\alpha-1}} \frac{e^{-u} \Theta du}{1-s\Theta}$$

$$= \frac{\cancel{\Theta^\alpha}}{\cancel{\Gamma(\alpha)}\cancel{\Theta^\alpha}} \cdot \frac{\cancel{\Gamma(\alpha)}}{(1-s\Theta)^\alpha} = \frac{1}{(1-s\Theta)^\alpha} \quad \text{if } s < \frac{1}{\Theta}$$

QED (2)

$$\textcircled{3} \quad \text{Say } X_k \sim \mathcal{P}(\alpha_k, \Theta) \quad \therefore \mathcal{M}_{X_k}(s) \stackrel{(2)}{=} \frac{1}{(1-s\Theta)^{\alpha_k}} \quad \text{if } s < \frac{1}{\Theta}$$

$$\therefore \mathcal{M}_x(s) = E[e^{sX}] = E\left[e^{s \sum_{k=1}^n X_k}\right] = E\left[e^{\sum_{k=1}^n sX_k}\right]$$

$$\begin{aligned}
 &= E \left[\prod_{k=1}^n e^{sX_k} \right] \stackrel{\text{ind}}{=} \prod_{k=1}^n E[e^{sX_k}] = \prod_{k=1}^n m_{X_k}(s) \\
 &\stackrel{\gamma}{=} \prod_{k=1}^n (1-s\theta)^{-\alpha_k} = (1-s\theta)^{-\sum_{k=1}^n \alpha_k} \\
 &\longleftrightarrow X \sim \gamma \left(\sum_{k=1}^n \alpha_k, \theta \right) \quad \text{QED (3)}
 \end{aligned}$$

CLT and χ^2

Say $X \sim \chi^2(r)$ $\therefore X = \sum_{k=1}^r X_k$ with $\begin{cases} \textcircled{1} \text{ independent } X_1, \dots, X_n \\ \textcircled{2} X_k \sim \chi^2(1) \text{ all } k \end{cases}$

$$\therefore \mu_{X_k} = 1 \quad \text{and} \quad \sigma_{X_k}^2 = 2 \quad (\text{since d.f.} = 1)$$

CLT: iid $\sum_{k=1}^r X_k \stackrel{\downarrow}{\approx} N(r\mu, r\sigma^2)$ for "large" r

$$\begin{aligned} \therefore X &\stackrel{\downarrow}{\approx} N(r\mu, r\sigma^2) \text{ if } X \sim \chi^2(r) \text{ and large } r \\ &= N(r \cdot 1, r \cdot 2) \end{aligned}$$

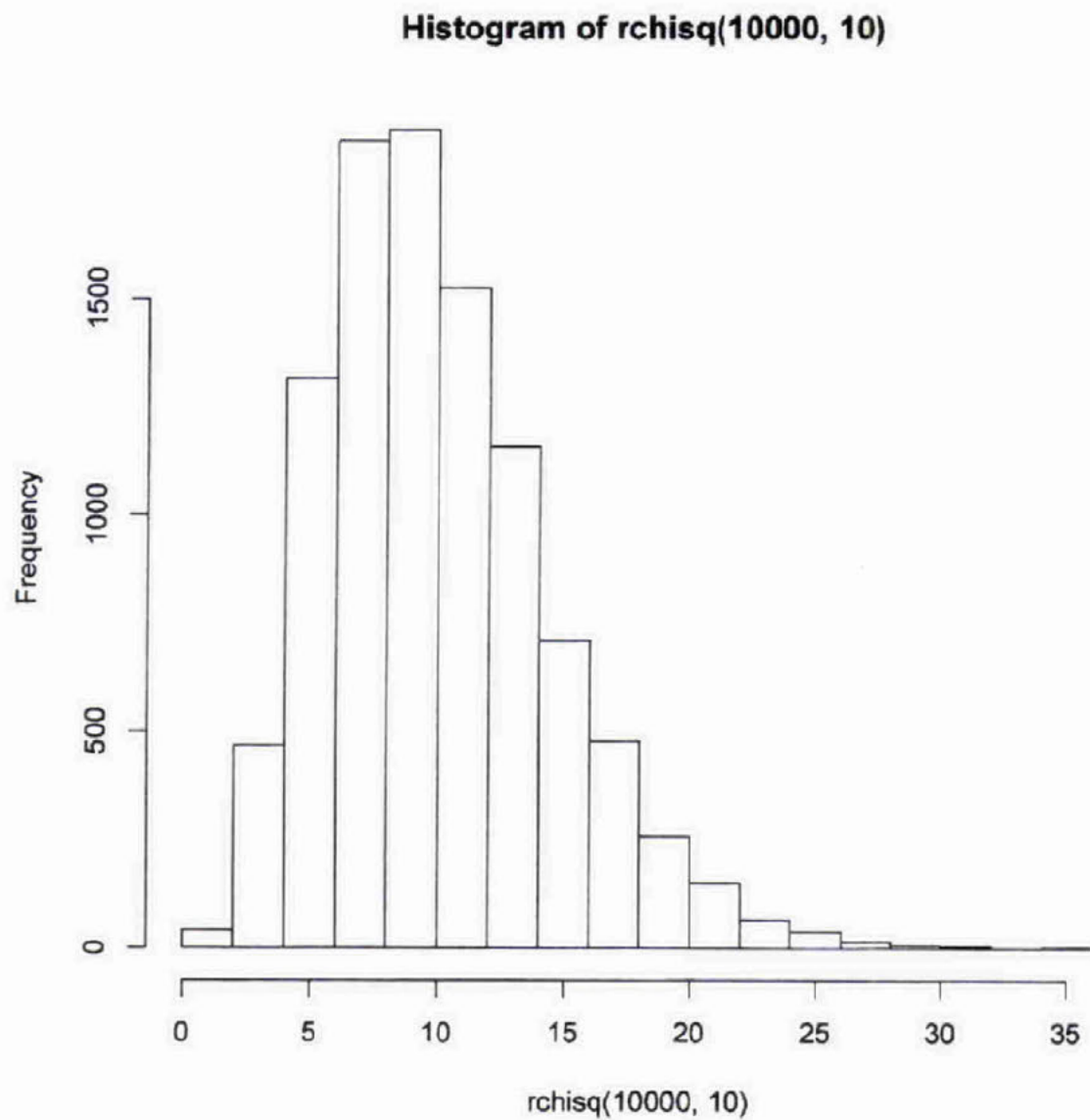
$$\therefore X \stackrel{\downarrow}{\approx} N(r, 2r)$$

Similarly: $\frac{1}{r} X = \frac{1}{r} \sum_{k=1}^r X_k = \overline{X}_r \stackrel{\text{CLT}}{\approx} N(1, \frac{2}{r})$

Ex: (in R) 10,000 iid realizations of $X \sim \chi^2(r)$

Compare: $\text{hist}(\text{rchisq}(10000, 10))$ $\swarrow r=10$

$\text{hist}(\text{rchisq}(10000, 100))$ $\swarrow r=100$

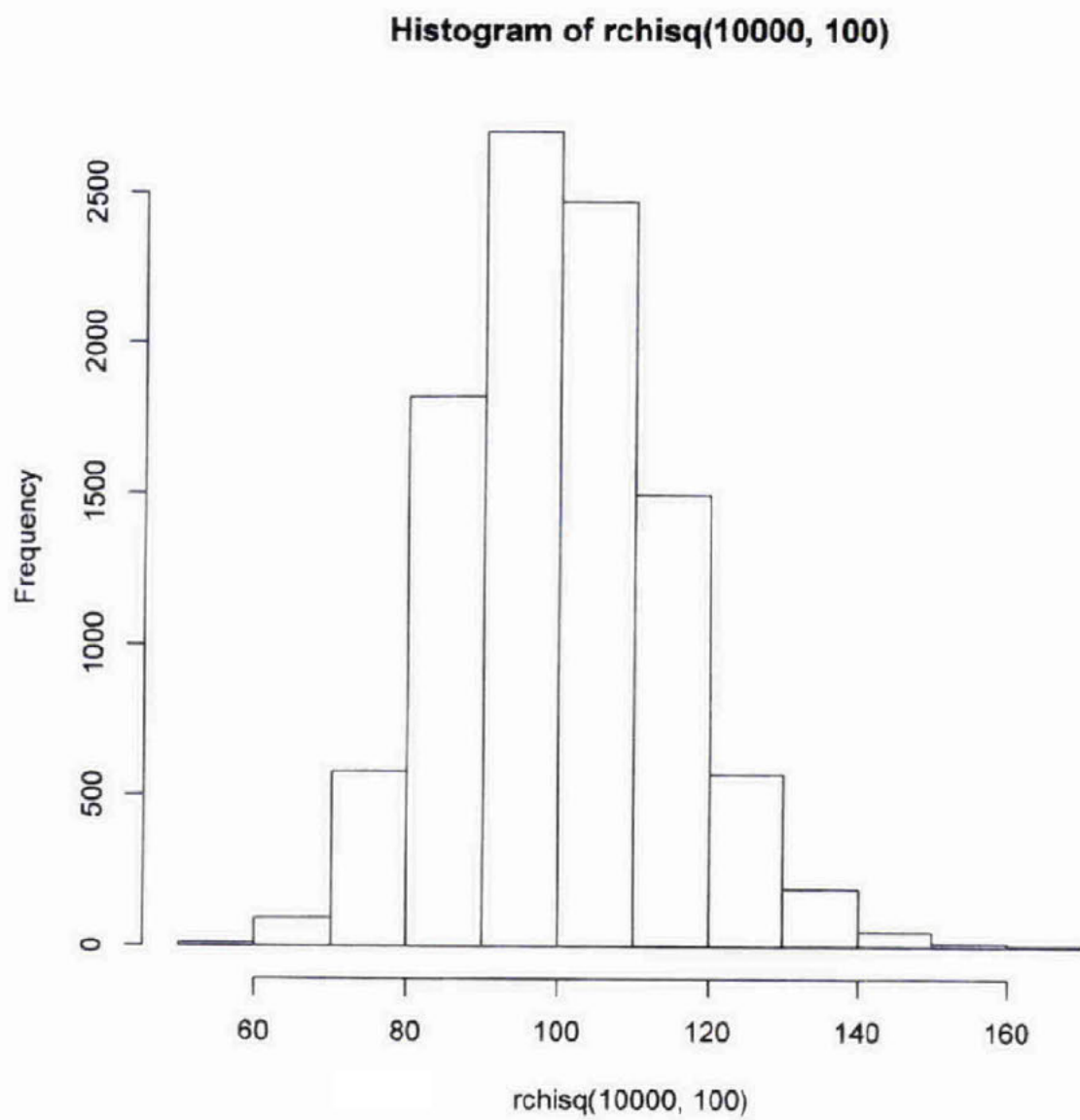


10,000 iid samples from $X \sim \chi^2(r)$

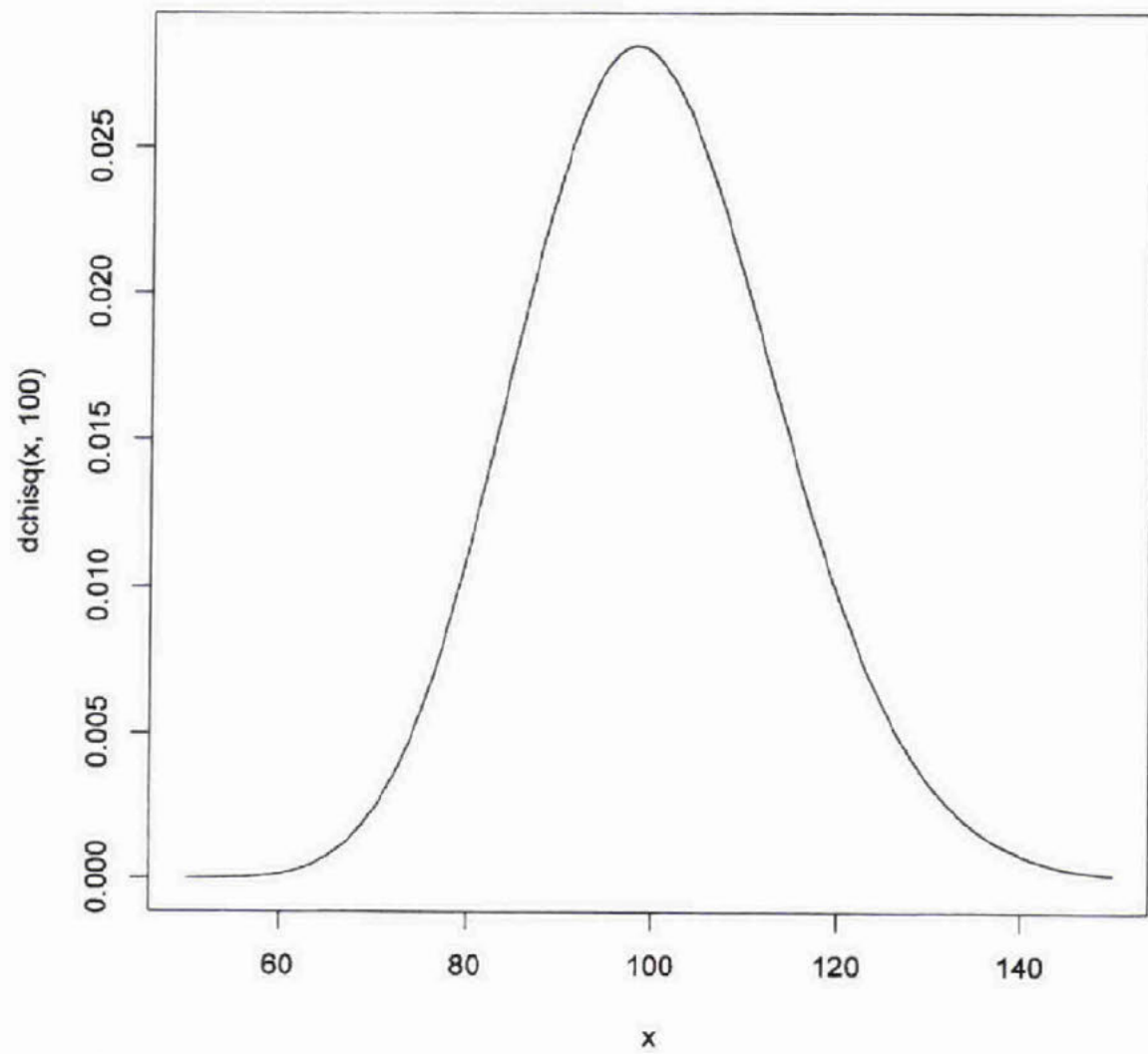
R command: `hist(rchisq(10000, 10))`

↑
 $n = \#$ random samples

↙ $r = 10$ d.f.

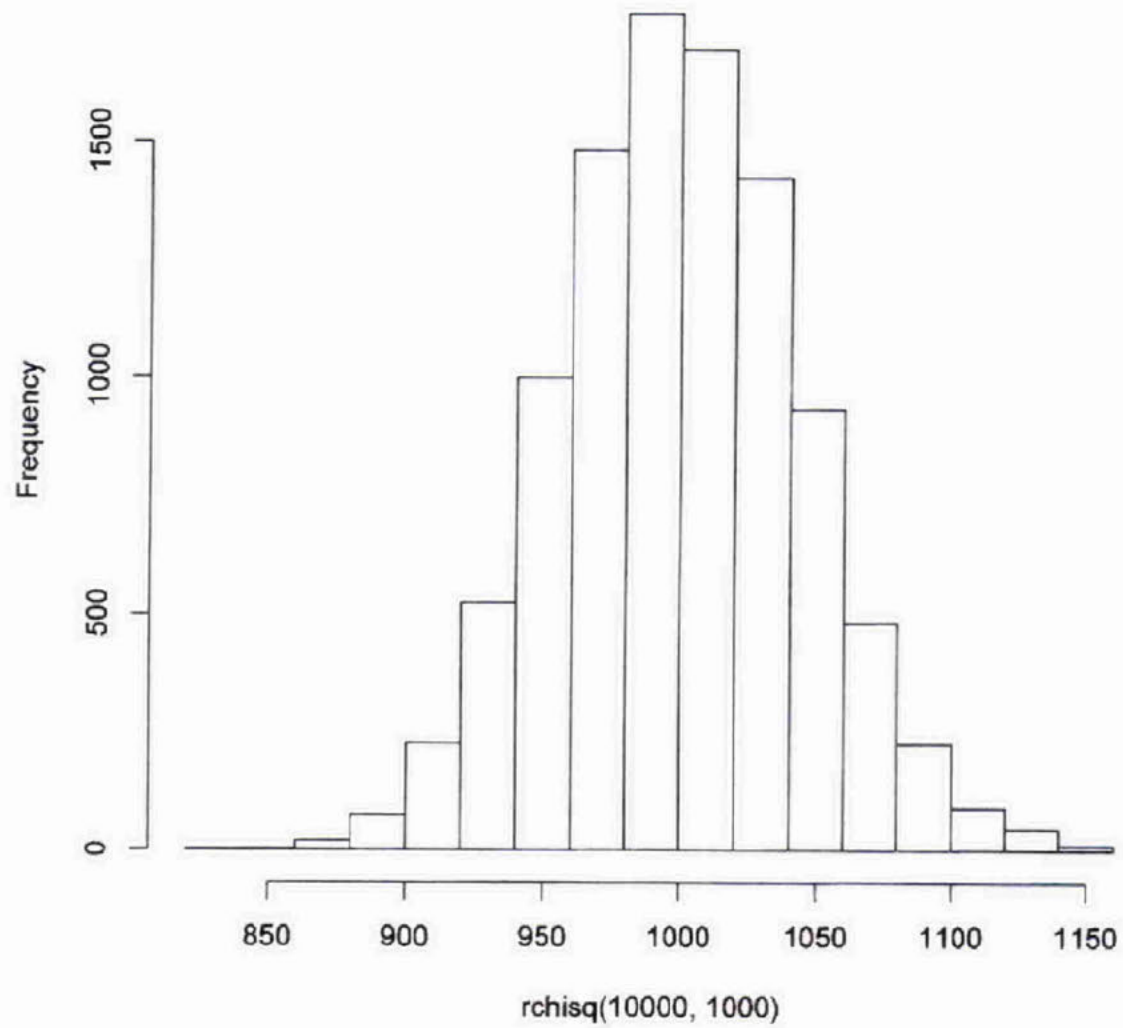


10,000 iid samples from $X \sim \chi^2(100)$

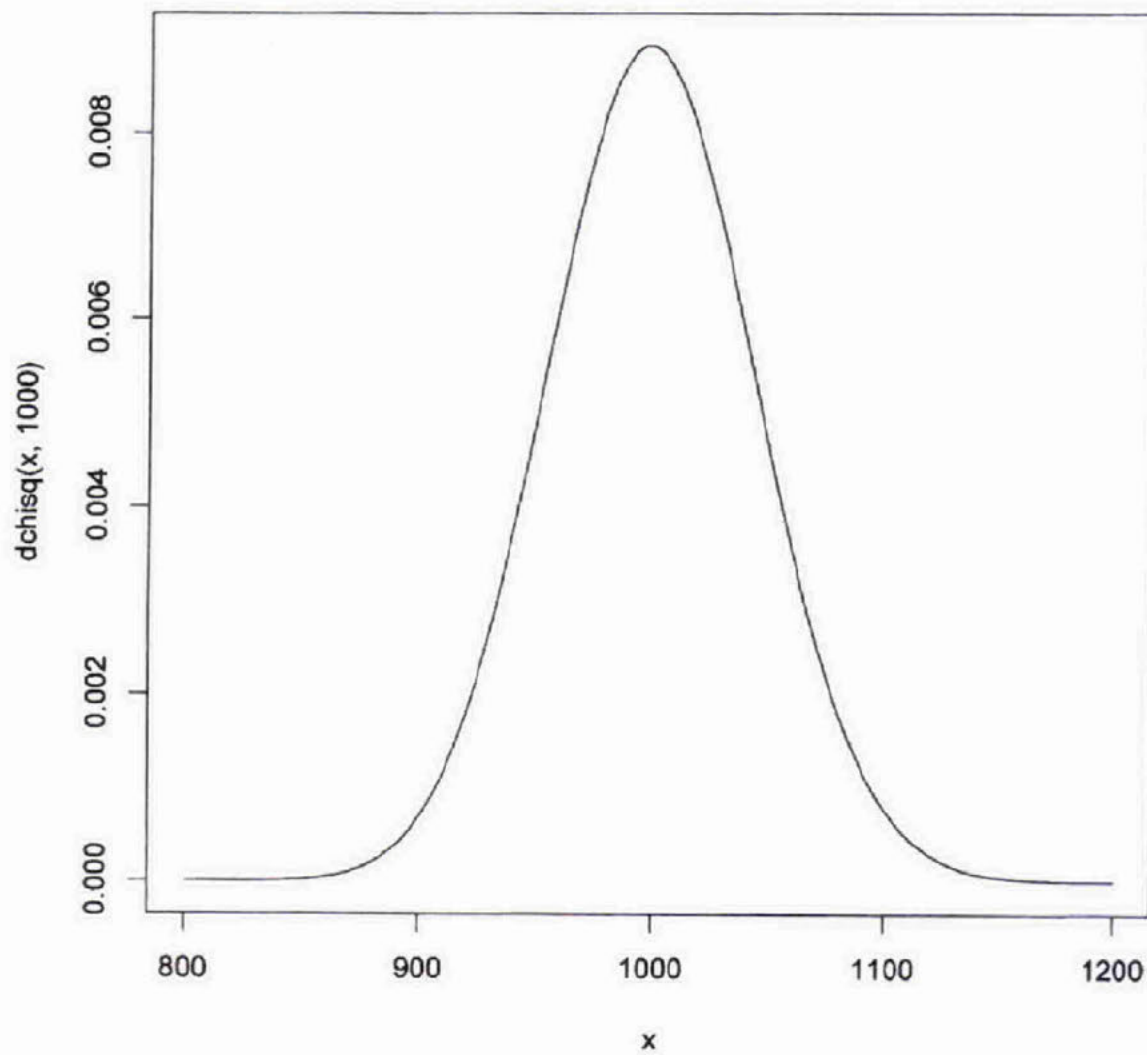


pdf of $X \sim \chi^2(100)$

Histogram of rchisq(10000, 1000)



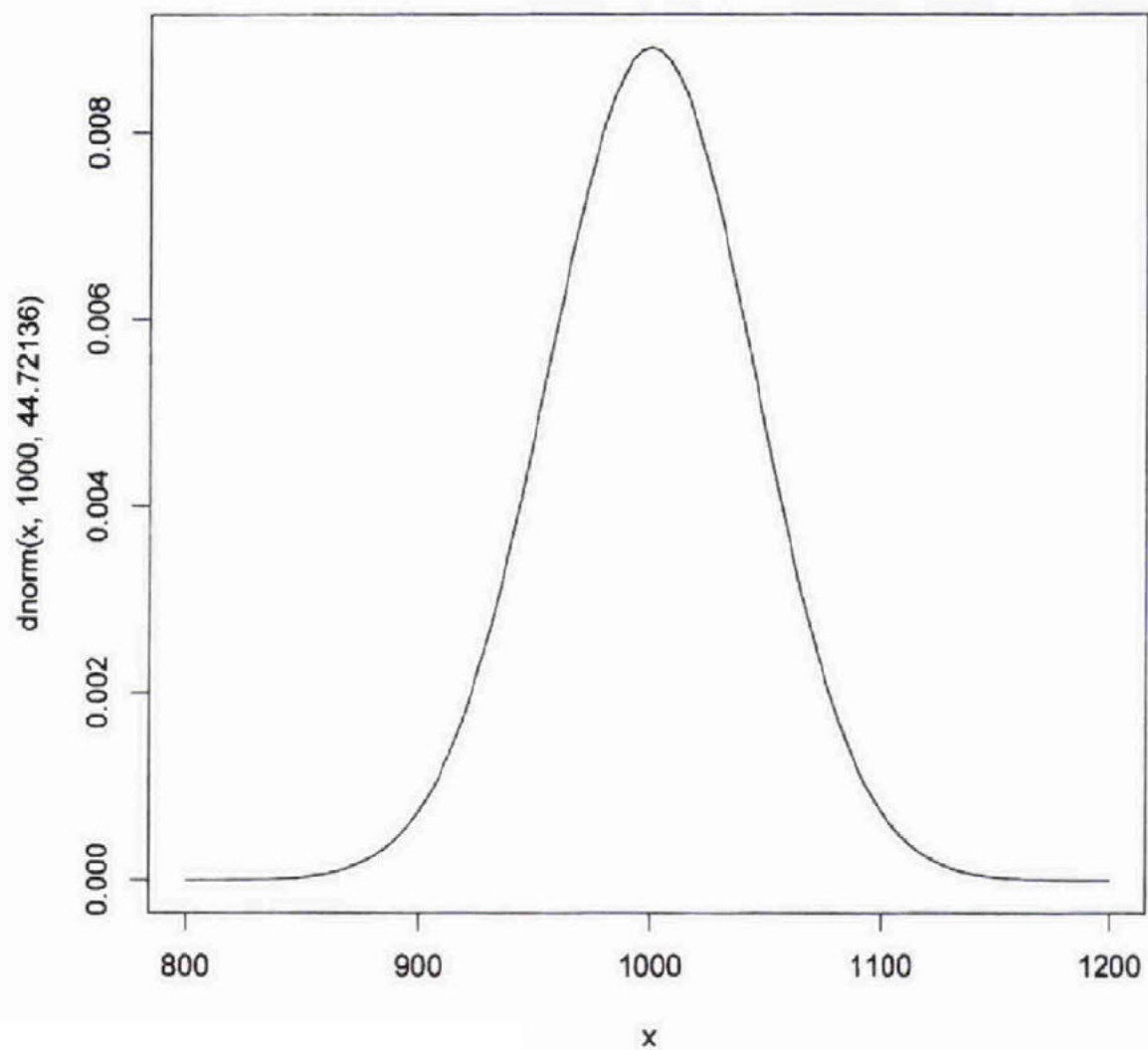
10,000 iid samples from $X \sim \chi^2(1000)$



pdf of $X \sim \chi^2(1000)$

$$\therefore X = \sum_{k=1}^{1000} X_k \stackrel{\text{CLT}}{\approx} N(1000, 2000)$$

for iid $X_k \sim \chi^2(1)$



pdf of $X \sim N(1000, 2000)$

$$\therefore \sigma_x = \sqrt{2000} = 44.72136$$

$\therefore$ Use in R: `curve(dnorm(x, 1000, 44.72136), 800, 1200)`