Lec 29 Math 241 11/29/16

Recap: $g(x)$ special mean... Recall the Taylor series expansion:

$M_X(t) := E[e^{tX}]$ def of mgf $\qquad e^{tX} = 1 + tX + \frac{t^2X^2}{2!} + \frac{t^3X^3}{3!} + \ldots$

Properties

(I) $M_X(t) = M_Y(t) \Rightarrow X \overset{d}{=} Y$

(II) $E[X^k] = M_X^{(k)}(0)$

(III) if $Y = aX + c \Rightarrow M_Y(t) = e^{tc} M_X(at)$

(IV) if X, Y indep $\qquad$ if X, Y iid

$M_{X+Y}(t) = M_X(t)\, M_Y(t) = (M_X(t))^2$

(V) Levy's Continuity thm

if $\lim_{n\to\infty} M_{X_n}(t) = M_Y(t) \iff X \longrightarrow Y$ the r.v. X converges to the r.v. Y

$\Rightarrow$ if n large $X_n \overset{d}{\approx} Y$ $\quad$ X, Y are approx. equally distr.

$\Rightarrow f_{X_n} \approx f_Y$ PDFs approx

---

if $X \sim \text{Bern}(p) \Rightarrow M_X(t) = 1 - p + pe^t$ (from def)

$X \sim \text{Binom}(n, p) \Rightarrow M_X(t) = (1 - p + pe^t)^n$ (rule IV)

$X \sim \text{Geom}(p) \Rightarrow M_X(t) = \frac{pe^t}{1 - e^t(1-p)}$ if $t < \ln\left(\frac{1}{1-p}\right)$ (from def)

$X \sim \text{Exp}(\lambda) \Rightarrow M_X(t) = \frac{\lambda}{\lambda - t}$ if $t < \lambda$ (from def)

$Z \sim N(0,1) \Rightarrow M_Z(t) = e^{\frac{t^2}{2}}$

$X \sim N(\mu, \sigma^2) \Rightarrow M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$ $\qquad X \sim \text{Deg}(c) \Rightarrow M_X(t) = e^{ct}$

"A weak"

LLN if $X_1, \dots, X_n \overset{iid}{\sim}$ something s.t. $E(X) < \infty$,

$\underbrace{\overline{X}_n}_{\text{r.v. sequence}} \to \underbrace{\mu}_{\text{another r.v.}}$

$Y \sim Deg(\mu)$

the r.v. only with value μ

Via Levy's Cont Thm...

$\lim_{n\to\infty} M_{\overline{X}_n}(t) = M_Y(t) = e^{t\mu}$

$$M_{\overline{X}_n}(t) \underset{\text{by def of } \overline{X}}{=} M_{\frac{X_1+\dots+X_n}{n}}(t) \underset{\text{(by III)}}{=} M_{X_1+\dots+X_n}\left(\frac{t}{n}\right) \underset{\text{(by IV)}}{=} \left(M_X\left(\frac{t}{n}\right)\right)^n \underset{\text{by def of mgf}}{=} \left(E\left[e^{\frac{t}{n}X}\right]\right)^n \underset{\text{by Taylor series}}{=} \left(E\left[1 + \frac{tX}{n} + \frac{t^2X^2}{n^2} + \frac{t^3X^3}{n^3} + \dots\right]\right)^n$$

We say $f(n) = o(g(n))$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$ i.e. $f(n)$ goes to ∞ slower than $g(n)$

e.g. $n^2 = o(n^3)$ since $\lim_{n\to\infty} \frac{n^2}{n^3} = \lim_{n\to\infty} \frac{1}{n} = 0$ ✓

We can say

$$\frac{t^2X^2}{n^2} + \frac{t^3X^3}{n^3} + \dots = o\left(\frac{1}{n}\right) \quad \text{why?} \quad \lim_{n\to\infty} \frac{\frac{t^2X^2}{n^2} + \frac{t^3X^3}{n^3} + \dots}{\frac{1}{n}} = \lim_{n\to\infty} \frac{t^2X^2}{n} + \frac{t^3X^3}{n^2} + \dots = 0 \; ✓$$

$$\Rightarrow M_{\overline{X}_n}(t) = \left(E\left[1 + \frac{tX}{n} + o\left(\frac{1}{n}\right)\right]\right)^n = \left(1 + \frac{t\mu}{n} + E\left[o\left(\frac{1}{n}\right)\right]\right)^n = \left(1 + \frac{t\mu}{n} + o\left(\frac{1}{n}\right)\right)^n$$

ignore this expectation as a technical pt.

$$\Rightarrow \lim_{n\to\infty} M_{\bar{X}_n}(t) = \lim_{n\to\infty}\left(1+\frac{t\mu}{n}+o\left(\frac{1}{n}\right)\right)^n$$

Recall $\lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n = e^a$

Also show $\lim_{n\to\infty}\left(1+\frac{a}{n}+o\left(\frac{1}{n}\right)\right)^n = e^a$ ← Thm (Symposium) but goes slow...

e.g. $\lim_{n\to\infty}\left(1+\frac{1}{n}+\frac{1}{n^{1.1}}\right)^n = e$ but if $n = 1$ billion $\Rightarrow f(n) = 3.08 \neq 2.718$

$n = 1$ million $\Rightarrow f(n) = 2.80 \neq 2.718$

Slow!

$$\Rightarrow \lim_{n\to\infty} M_{\bar{X}_n}(t) = \underbrace{e^{t\mu}}_{\text{the mgf for Deg}(\mu)} \quad \text{thus} \quad \bar{X} \to \mu$$

~~$\Rightarrow \lim_{n\to\infty} M_{\bar{X}}(t) = e^{t\mu} = M_{\mu}(t) \Rightarrow \bar{X} \to \mu$ proof of LLN.~~

Still... why is $Z \sim N(0,1)$ so special?

$X_1, \dots, X_n \overset{iid}{\sim}$ Everything with mean μ and s.e. σ

Consider
$$C_n := \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$$
C_n is the standardized $\bar{X}_n$

$$\left.\begin{aligned} E(C_n) &= 0 \\ SE(C_n) &= 1 \end{aligned}\right\} \text{why?}$$

Prestep:

$$\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}$$
$$= \frac{\sqrt{n}\left(\frac{X_1 + \dots + X_n}{n}\right) - \mu}{\sigma} \qquad \mu = \frac{n\mu}{n}$$
$$= \frac{\sqrt{n}\ \frac{X_1 + \dots + X_n}{n} - \frac{\mu + \dots + \mu}{n}}{\sigma}$$
$$= \frac{(X_1 - \mu) + \dots + (X_n - \mu)}{\sigma\sqrt{n}}$$
$$= \frac{1}{\sqrt{n}}\left(\frac{X_1 - \mu}{\sigma} + \dots \frac{X_n - \mu}{\sigma}\right) \qquad \text{let } Z_i := \frac{X_i - \mu}{\sigma}$$
$$= \frac{1}{\sqrt{n}}\left(Z_1 + \dots + Z_n\right)$$

Note: $E(Z_i) = 0$, $SE(Z_i) = 1$ why?

$$M_{C_n}(t) = M_{\frac{1}{\sqrt{n}}(Z_1 + \dots + Z_n)}(t) = M_{Z_1 + \dots + Z_n}\left(\frac{t}{\sqrt{n}}\right) = \left(M_Z\left(\frac{t}{\sqrt{n}}\right)\right)^n$$

$$= \left(E\left[1 + \frac{tZ}{\sqrt{n}} + \frac{t^2Z^2}{2n} + \frac{t^3Z^3}{3!\,n^{1.5}} + \frac{t^4Z^4}{4!\,n^2} + \dots\right]\right)^n$$

since $\mu = 0$ $\qquad$ $\frac{t^2}{n}$ since $\sigma = 1$

$$= \left(1 + \cancelto{0}{\frac{t\mu}{\sqrt{n}}} + \cancel{\frac{t^2\sigma^2}{2n}} + E\underbrace{\left[\frac{t^3Z^3}{3!\,n^{1.5}} + \frac{t^4Z^4}{4!\,n^2} + \dots\right]}_{o\left(\frac{1}{n}\right)}\right)^n$$

ignore $E\left(o\left(\frac{1}{n}\right)\right)$ $\qquad$ same as before

$$= \left(1 + \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)^n$$

What happens as $n \to \infty$

$$\lim_{n\to\infty} M_{C_n}(t) = \lim_{n\to\infty}\left(1 + \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)^n = e^{\frac{t^2}{2}} = M_Z(t) \text{ the std. normal!}$$

$\Rightarrow C_n \longrightarrow Z \sim N(0,1)$ Central limit theorem! (CLT)

CLT if $X_1, \dots, X_n \overset{iid}{\sim}$ with mean μ, s.d. σ, then....

$$\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \longrightarrow N(0,1) \qquad (\mathrm{I})$$

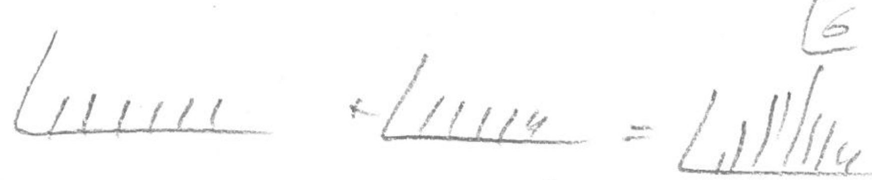
Corollaries....

$$\Rightarrow \text{if } n \text{ large} \quad \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \overset{d}{\approx} N(0,1) \qquad (\mathrm{II})$$

$$\Rightarrow \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \approx Z \Rightarrow \overline{X} \approx \frac{\sigma}{\sqrt{n}} Z + \mu \Rightarrow \overline{X} \overset{d}{\approx} N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right) \qquad (\mathrm{III})$$

$$\Rightarrow n\overline{X} \approx \sigma\sqrt{n}\,Z + n\mu \Rightarrow T \approx \sigma\sqrt{n}\,Z + n\mu \Rightarrow T \overset{d}{\approx} N\left(n\mu, (\sigma\sqrt{n})^2\right) \qquad (\mathrm{IV})$$

$X \sim \text{Bathtub}(\mu)$



Already it begins!

$X_1 + \dots + X_n \approx N(n\mu, (\sqrt{n}\sigma)^2)$

$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \approx e^{-x^2}$ makes tails small!

natural balance between means & errors

When is a certain value considered "normal"?

Why is it called the normal distr??

Example $X_1, \dots, X_{30} \overset{iid}{\sim} \text{Geom}(\frac{1}{2})$

What's the prob the avg wait time is more than 2.75?

standardize!!

$$P(\overline{X} \geq 2.75) \approx P\left(\frac{\overline{X} - 0.5}{.258} \geq \frac{2.75 - 0.5}{.258}\right) = P(Z \geq 3) = .0015$$

Recall $\overline{X} \approx N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right) \Rightarrow \overline{X} \approx N(0.5, .258^2)$

$\mu = p = \frac{1}{2}, \ \sigma = \sqrt{\frac{1-p}{p^2}} = \sqrt{\frac{1/2}{1/4}} = \sqrt{2} \approx 1.414 \Rightarrow \frac{\sigma}{\sqrt{n}} \approx \frac{1.414}{\sqrt{30}} \approx .258$

Random walk.. take 100 steps, what's the prob we are more than 10 steps from where we started?

$X_1,\dots,X_{100} \overset{iid}{\sim}$ Rademacher

$T = X_1+\dots+X_{100}$

total distance away

$$P(|T|\ge 10) = P(T>10)+P(T<-10)$$
$$\approx P\left(\frac{T-0}{10} > \frac{10-0}{10}\right) + P\left(\frac{T-0}{10} < \frac{-10-0}{10}\right) = P(Z>1)+P(Z<-1) = 2P(Z\ge 1)$$

symmetric bell curve

$$= 2\cdot 16\% = \boxed{32\%}$$

$$T \overset{d}{\approx} N(n\mu, (\sqrt{n}\sigma)^2) = N(0, 10^2)$$

$\mu = 0, \ \sigma = 1$

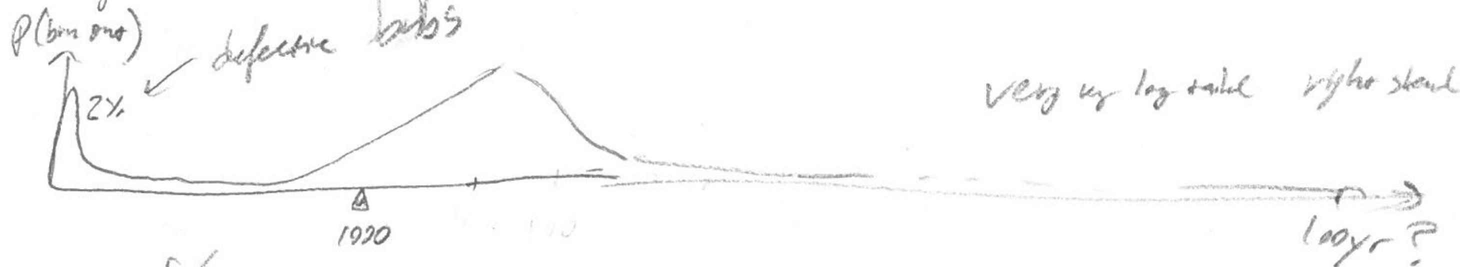
$E(X)$ ... $\sqrt{Var(X)}$

previous lectures

$\Rightarrow n\mu = 100\cdot 0 = 0$

$\sqrt{n}\sigma = \sqrt{100}\cdot 1 = 10$

Light failure



$$X \sim f(\mu \approx 1000, \ \sigma \approx 500)$$

You get 50 bulbs. What's the prob the avg burnout is > 1300hr?

$X_1,\dots,X_{50} \overset{iid}{\sim} f$

$$P(\bar{X} > 1300hr) \approx P\left(\frac{\bar{X}-1000}{70.7} > \frac{1300-1000}{70.7}\right) \approx P(Z > 4.24) \approx 0$$

$$\bar{X} \approx N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right) = N(1000, 70.7^2)$$

$$\frac{500}{\sqrt{50}} = 70.7$$