Lioneta Osorio
Math 241
November 24, 2015 — Lecture 20

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

$X_1, \dots, X_n \overset{iid}{\sim}$ something with $\mu, \sigma^2 < \infty$

$C_n := \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} = \dots = \frac{Z_1}{\sqrt{n}} + \dots + \frac{Z_n}{\sqrt{n}}$ $\qquad Z_i := \frac{X-\mu}{\sigma}$

Goal: How is C_n distributed as n gets larger?

Moment generating Function

Recall $M_X(t) = E[e^{tX}] = \underbrace{E[1 + tX + \frac{t^2X^2}{2!} + \dots]}_{\text{Taylor Expansion}}$

$\lim_{n\to\infty} M_{C_n}(t) = \left(M_{\frac{Z}{\sqrt{n}}}(t)\right)^n = \left(M_Z\left(\frac{t}{\sqrt{n}}\right)\right)^n$

$= \left(1 + \frac{t}{\sqrt{n}}E(Z) + \frac{t^2}{n}\frac{E(Z^2)}{2} + \underbrace{\frac{t^3}{n^{3/2}}\frac{E(Z^3)}{3!} + \frac{t^4}{n^2}\frac{E(Z^4)}{4!} + \dots}_{tail(n)}\right)^n$

$Z_i := \frac{X-\mu}{\sigma} \longrightarrow E(Z) = 0,\ E(Z^2) = 1$

$\lim_{n\to\infty}\left(1 + \frac{t^2/2}{n} + tail(n)\right)^n$ $\qquad e = \lim_{n\to\infty}\left(1 + \frac{1}{n} + O\left(\frac{1}{n}\right)\right)^n$

$= \lim_{n\to\infty}\left(1 + \frac{t^2/2}{n} + o\left(\frac{1}{n}\right)\right)^n = e^{t^2/2}$

Is $tail(n)$ an $(\frac{1}{n})(0)$ function?

$\lim_{n\to\infty} \frac{\frac{t^3}{n^{3/2}}\frac{E(Z^2)}{3!} + \frac{t^4}{n^2}\frac{(EZ^4)}{4!} \dots}{\frac{1}{n}} = \lim_{n\to\infty} \frac{t^3}{\sqrt{n}}\frac{E(Z^3)}{3!} + \frac{t^4}{n}\frac{E(Z^4)}{4!} + \dots = 0$

Central Limit theorem (CLT)

$\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$ if n is large $\Rightarrow \bar{X} = \frac{\sigma}{\sqrt{n}}C_n + \mu \sim N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right)$

$T = n\bar{X} \sim N(n\mu, (\sqrt{n}\,\sigma)^2)$

$X_1, \dots, X_{30} \overset{iid}{\sim} Geom(\frac{1}{2})$ What is the probability I wait more than 2.75 sec on average?

Probability Statement: $P(\bar{X} > 2.75)$

$\bar{X} \sim N\left(\mu, \left(\frac{\sigma}{\sqrt{n}}\right)^2\right)$ by CLT $\to N(2, .258^2)$

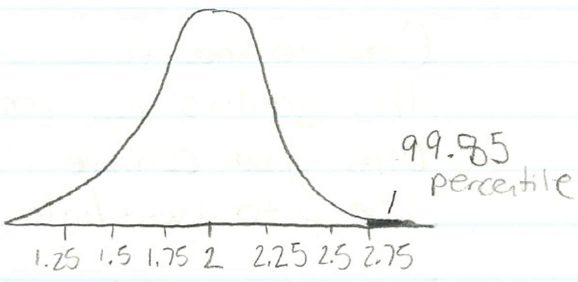
$\mu = \frac{1}{p} = \frac{1}{\frac{1}{2}} = 2$

$\sigma = \sqrt{\frac{1-p}{p^2}} = \sqrt{\frac{1-\frac{1}{2}}{(\frac{1}{2})^2}} = \sqrt{2}$

$P(\bar{X} > 2.75) =$

$= P\left(\frac{\bar{X}-2}{.258} > \frac{2.75-2}{.258}\right)$

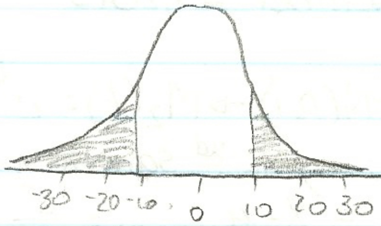
$= P(Z > 3) = .0015$



Random Walk of 100 steps

What is the probability you are more than 10 steps from the starting position?

$X_1, \dots, X_{100} \overset{iid}{\sim} \begin{cases} +1 & wp \ \frac{1}{2} \\ -1 & wp \ \frac{1}{2} \end{cases}$ — $\mu = 0$, $\sigma = 1$



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

$\overset{\wedge}{\text{by CLT}}$

$P(|T| > 10) = 2P(T > 10) = 2P\left(\frac{T-0}{10} > \frac{10-0}{10}\right) = 2P(Z > 1) = 2 \cdot .16 = \boxed{.32}$

Shipments are late 2% of the time. What is the probability that more than 3% are late on average among 10,000 shipments? Assume iid

$X \sim \text{Bernoulli}(.02)$

$X_1, \dots, X_{10000} \overset{iid}{\sim} \text{Bern}(.02)$ $\mu = .02$ $\frac{\sigma}{\sqrt{n}} = \sqrt{\frac{p(1-p)}{n}} = \frac{\sqrt{.02 \cdot .98}}{10000} = .0014$

$P(\bar{X} > .03)$ $\bar{X} \sim N(\mu, (\frac{\sigma}{\sqrt{n}})^2)$

$\overset{\wedge}{\text{by CLT}}$

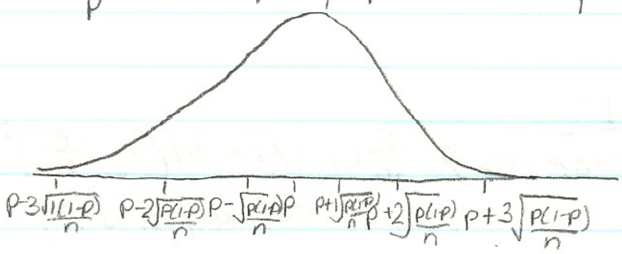
$P\left(\frac{\bar{X} - .02}{.0017} > \frac{.03 - .02}{.0014}\right) = P(Z > 7.14) \approx \boxed{0}$

$X_1, \dots, X_n \overset{iid}{\sim} \text{Bern}(p)$

$\bar{X}$ $\bar{x} = \frac{\Sigma 1\text{'s}}{n}$

$\uparrow$ $\uparrow$

$\hat{p}$ $\hat{p}$ → "sample proportion" $\hat{P} \sim N\left(p, \left(\sqrt{\frac{p(1-p)}{n}}\right)^2\right)$



$\hat{p} = \frac{6}{14} = .43$ $X_1, \dots, X_{14} \overset{iid}{\sim} \text{Bern}(p)$

what is this?

① parameter of model

② "population proportion" *

③ true proportion of mushroom liking people

Goal: to know p

this goal is <u>not</u> possible

Other goal: estimate p

Use $\hat{p}$ to guess/infer p

Point Estimate:

One point guess

$p \approx \hat{p}$ because $\bar{x} \approx \mu$ since LLN

Interval Estimate:

Range of values for p.