Class 23: LLN, CLT, and final review

Thursday, December 5, 2019 11:38 AM

- HW 8 due Fri, Dec 6, <u>11pm</u>

Let $X_1, X_2, \ldots$ be iid with mean $m = \mathbb{E}[X_i]$
variance $\sigma^2 = \operatorname{Var}(X_i) > 0$.

Define the sample average $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

<u>LLN</u>: $\bar{X}_n \to m$ with prob. 1 as $n \to \infty$,
or $\bar{X}_n \approx m$ for n large.

<u>CLT</u>: $\bar{X}_n \approx m + \frac{\sigma}{\sqrt{n}} Z$ where $Z \sim N(0,1)$.
standard normal

<u>Idea of proof:</u>

- Inversion theorem for moment generating functions (MGFs) suggests that it's enough to show that MGF of

$$Z_n = \sqrt{n}\left(\frac{\bar{X}_n - m}{\sigma}\right)$$ converges to the MGF of Z.

- The MGF of $Z \sim N(0,1)$ we saw to be $M_Z(s) = e^{\frac{1}{2}s^2}$.
- Let $M_X(s) = \mathbb{E}[e^{sX_i}]$ be MGF of X_i's.

Then, for $s \in \mathbb{R}$,
$$M_{Z_n}(s) = \mathbb{E}[e^{sZ_n}] = \mathbb{E}\left[e^{s\sqrt{n}\left(\frac{\bar{X}_n - m}{\sigma}\right)}\right].$$

Rewrite
$$Z_n = \sqrt{n}\left(\frac{\bar{X}_n - m}{\sigma}\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\underbrace{\left(\frac{X_i - m}{\sigma}\right)}_{\text{normalized error}}.$$

Let $Y_i = \frac{X_i - m}{\sigma}$. Then $\mathbb{E}[Y_i] = \frac{\mathbb{E}[X_i] - m}{\sigma} = \frac{m - m}{\sigma} = 0$, and $\quad Var(Y_i) = Var\left(\frac{X_i - m}{\sigma}\right) = Var\left(\frac{X_i}{\sigma}\right) = \frac{1}{\sigma^2} Var(X_i)$

$$= \frac{\sigma^2}{\sigma^2} = 1.$$

Let $M_Y(s) = \mathbb{E}[e^{sY_i}]$ for $s \in \mathbb{R}$, MGF of Y_i's.

Then MGF of Z_n

$$M_{Z_n}(s) = \mathbb{E}[e^{sZ_n}] = \mathbb{E}\left[e^{\frac{s}{\sqrt{n}} \sum_{i=1}^{n} Y_i}\right]$$

$$= \mathbb{E}\left[e^{\frac{s}{\sqrt{n}} Y_1} e^{\frac{s}{\sqrt{n}} Y_2} \cdots e^{\frac{s}{\sqrt{n}} Y_n}\right].$$

The Y_i's are independent because the X_i's are.

So

$$M_{Z_n}(s) = \mathbb{E}\left[e^{\frac{s}{\sqrt{n}} Y_1}\right] \mathbb{E}\left[e^{\frac{s}{\sqrt{n}} Y_2}\right] \cdots \mathbb{E}\left[e^{\frac{s}{\sqrt{n}} Y_n}\right]$$

$$= M_Y\left(\tfrac{s}{\sqrt{n}}\right) M_Y\left(\tfrac{s}{\sqrt{n}}\right) \cdots M_Y\left(\tfrac{s}{\sqrt{n}}\right)$$

$$= \left(M_Y\left(\tfrac{s}{\sqrt{n}}\right)\right)^n.$$

Use Taylor approximation around $s = 0$:

<u>Note</u>: $M_Y(0) = \mathbb{E}[e^{0 \cdot Y_i}] = 1$

$M_Y'(0) = \mathbb{E}[Y_i] = 0$

$M_Y''(0) = \mathbb{E}[Y_i^2] = Var(Y_i) = 1.$ since $\mathbb{E}[Y_i] = 0$

Thus $M_Y(h) \approx M_Y(0) + h\cancelto{0}{M_Y'(0)} + \frac{1}{2}h^2 M_Y''(0)$ for $h > 0$ small.

$$= 1 + \tfrac{1}{2}h^2.$$

Use $h = s/\sqrt{n}$ to get

$$M_{Z_n}(s) = \left(M_Y(s/\sqrt{n})\right)^n$$

$$\approx \left(1 + \frac{s^2}{2n}\right)^n$$

$$\to e^{s^2/2} \quad \text{as } n \to \infty.$$

Recall:

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

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

use $x = s^2/2$

Summary: MGF of Z_n converges to MGF of $Z \sim N(0,1)$.

Similar proof (based on MGF) for LLN:

The MGF of $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ is:

$$M_{\bar{X}_n}(s) = \mathbb{E}[e^{s\bar{X}_n}] = \mathbb{E}[e^{\frac{s}{n}\sum_{i=1}^{n} X_i}]$$
$$= \mathbb{E}[e^{\frac{s}{n}X_1} e^{\frac{s}{n}X_2} \cdots e^{\frac{s}{n}X_n}]$$
$$= \mathbb{E}[e^{\frac{s}{n}X_1}]\mathbb{E}[e^{\frac{s}{n}X_2}] \cdots \mathbb{E}[e^{\frac{s}{n}X_n}] \qquad \text{by independence}$$
$$= \left(M_X(\tfrac{s}{n})\right)^n.$$

Use 1st order Taylor approximation:

$$M_X(h) \simeq M_X(0) + hM_X'(0) \qquad M_X'(0) = \mathbb{E}[X_i] = m$$
$$= 1 + hm, \qquad h > 0 \text{ small.}$$

Use $h = s/n$ to get

$$M_{\bar{X}_n}(s) \simeq \left(1 + \frac{sm}{n}\right)^n \longrightarrow e^{sm} \quad \text{as } n \to \infty.$$

Now e^{sm} is the MGF of the constant r.v. $= m$.

The MGF of $\bar{X}_n$ converges to MGF of the constant m,

and so $\bar{X}_n \to m$.

Applications of CLT:

Recall $X_1, X_2, \ldots$ iid mean m variance σ^2.

Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, and $Z_n = \sqrt{n}\left(\frac{\bar{X}_n - m}{\sigma}\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{X_i - m}{\sigma}\right)$.

CLT says $\bar{X}_n = m + \frac{\sigma}{\sqrt{n}} Z_n \simeq m + \frac{\sigma}{\sqrt{n}} Z$

or $Z_n \simeq Z$ where $Z \sim N(0,1)$.

To calculate probabilities involving $\bar{X}_n$, we can approximate by Z.

(Useful because distribution of $\bar{X}_n$ is usually hard to find.)

To calculate probabilities involving $\bar{X}_n$, we can approximate by Z.

(Useful because distribution of $\bar{X}_n$ is usually hard to find.)

<u>Ex:</u> For $a < b$, use $\bar{X}_n \simeq m + \frac{\sigma}{\sqrt{n}} Z$

$$\mathbb{P}(a \le \bar{X}_n \le b) \simeq \mathbb{P}\left(a \le m + \frac{\sigma}{\sqrt{n}} Z \le b\right)$$

$$= \mathbb{P}\left(\sqrt{n}\left(\frac{a-m}{\sigma}\right) \le Z \le \sqrt{n}\left(\frac{b-m}{\sigma}\right)\right).$$

$$= \dots$$

<u>Note:</u> CDF of $Z \sim N(0,1)$, $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$,

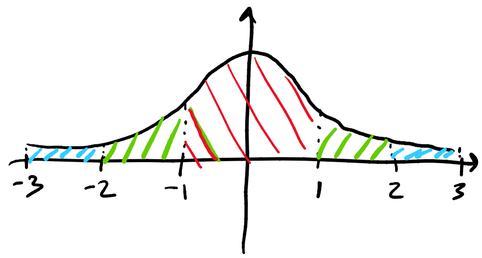
does not have a closed form.

Some particular values:

(1) $\mathbb{P}(|Z| \le 1) \simeq .68$

(2) $\mathbb{P}(|Z| \le 2) \simeq .95$

(3) $\mathbb{P}(|Z| \le 3) \simeq .998$



<u>Ex 1:</u> Flip n coins.

- How many coins must we flip in order to be 95% sure that the number of heads is within 5% of its mean?
- Let $X_i \sim Ber(\frac{1}{2})$ be iid. ($X_i = 1$ means heads, $X_i = 0$ tails.)

Then $\sum_{i=1}^{n} X_i$ = number of heads in n flips.

True mean is $\frac{n}{2} = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right]$.

- <u>Goal:</u> Find n such that

$\mathbb{P}$(being within 5% of mean) $\ge .95$.

Or $\mathbb{P}\left(\left|\sum_{i=1}^{n} X_i - \frac{n}{2}\right| \le .05n\right) \ge .95$.

or $\mathbb{P}\left(\left|\sum_{i=1}^{n} X_i - \frac{n}{2}\right| \leq .05n \right) \geq .95.$

or $\mathbb{P}\left(|\bar{X}_n - \frac{1}{2}| \leq .05 \right) \geq .95$ where $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$

Use $\bar{X}_n \simeq m + \frac{\sigma}{\sqrt{n}} Z$ where $Z \sim N(0,1)$

and $m = \mathbb{E}[X_i] = \frac{1}{2}$, $\sigma^2 = Var(X_i) = \frac{1}{4}$, $\sigma = \frac{1}{2}$.

So $\bar{X}_n \simeq \frac{1}{2} + \frac{1}{2\sqrt{n}} Z$.

Plug this in to get

$$\mathbb{P}\left(|\bar{X}_n - \tfrac{1}{2}| \leq .05 \right) \simeq \mathbb{P}\left(\left|\left(\tfrac{1}{2} + \tfrac{1}{2\sqrt{n}} Z\right) - \tfrac{1}{2}\right| \leq .05 \right)$$

$$= \mathbb{P}\left(|Z| \leq \frac{\sqrt{n}}{10} \right).$$

Want this $\leq .95$.

We know from above $\mathbb{P}(|Z| \leq 2) \simeq .95$.

So take n such that $\frac{\sqrt{n}}{10} = 2$,

or $n \geq 20^2 = 400$.

↳ 400 flips suffices.

Similarly, if we want 99.8% certainty, then we want

n such that $\mathbb{P}\left(|\bar{X}_n - \frac{1}{2}| \leq .05 \right) \geq .998$

or $\mathbb{P}\left(|Z| \leq \frac{\sqrt{n}}{10} \right) \geq .998.$

We know $\mathbb{P}(|Z| \leq 3) \simeq .998$

so we set $\frac{\sqrt{n}}{10} = 3$, or $n = 30^2 = 900$.

Final Exam Overview (cumulative!)

① Sample Spaces:

- working probabilities of events e.g. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ for A, B disjoint

- ... probability

- working probabilities of events e.g. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ for A, B disjoint
- conditional probability

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

↳ Bayes Rule, law of total probabilities

- independence of events $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$
- counting problems

② Random Variables

- PMFs / PDFs , CDFs , MGFs

↑ discrete ↑ continuous

$p(k) = \mathbb{P}(X = k)$ $f(x) = \frac{dF}{dx}(x)$ $F(x) = \mathbb{P}(X \le x)$ $M(s) = \mathbb{E}[e^{sX}]$

Key fact: $\frac{d^n M}{ds^n}(0) = \mathbb{E}[X^n]$

- Calculate probabilities, expectations, variance ...

Expectations:

$$\mathbb{E}[g(X)] = \begin{cases} \sum_k g(k) \underbrace{p(k)}_{\text{PMF of } X} & \text{if discrete} \\ \int_{-\infty}^{\infty} g(x) \underbrace{f(x)}_{\text{PDF of } X} dx & \text{if continuous} \end{cases}$$

Probabilities:

$$P(a < X \le b) = \underbrace{F}_{\text{CDF of } X}(b) - F(a) = \begin{cases} \sum_{a < k \le b} p(k) & \text{if discrete} \\ \int_a^b f(x)dx & \text{if continuous} \end{cases}$$

Variance:

$$\begin{aligned} Var(X) &= \mathbb{E}[(X - \mathbb{E}[X])^2] \\ &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \end{aligned}$$

- Transformation of RVs

Ex: If X is given and $Y = X^2 + 2X$,

find the PMF / PDF of Y using that of X.

continuous case, find CDF of Y first, then differentiate

- Common Distributions: (PMF / PDF, what do they model?)

Discrete

- Bernoulli, $Ber(p)$
  success/failure
- Binomial, $Bin(n,p)$
  - n indep. trials
  - sum of iid Bernoulli's
- Geometric, $Geom(p)$
  # trials until 1st success
- Poisson(λ), $\lambda > 0$
  rare events

Continuous

- Uniform, $Unif[a,b]$
- Exponential, $Exp(\lambda)$
  time until 1st arrival
- Normal, $N(0,1)$

(3) Multiple random variables (joint distributions)

- Joint PMFs & PDFs	$p_{X,Y}(n,k)$	$f_{X,Y}(x,y)$
- Marginals:	$p_X(n) = \sum_k p_{X,Y}(n,k)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$
- Conditionals:	$p_{X\|Y}(n\|k) = \frac{p_{X,Y}(n,k)}{p_Y(k)}$	$f_{X\|Y}(x\|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$
- Conditional Expectations:	$E[g(X)\|Y=y] = \sum_n g(n)\,p_{X\|Y}(n\|y)$	$= \int_{-\infty}^{\infty} g(x)\,f_{X\|Y}(x\|y)\,dx$

↳ as a random variable: $E[X|Y]$

↳ tower property: $\underbrace{E[X]}_{\text{would calculate using } f_X} = \overbrace{E[\underbrace{E[X|Y]}_{\text{calculate using } f_{X|Y}}]}^{\text{calculate using } f_Y}$

- Independence: $p_{X,Y}(x,y) = p_X(x)\,p_Y(y)$ (discrete) or $f_{X,Y}(x,y) = f_X(x)\,f_Y(y)$ (continuous)

Ex: If $X \sim Exp(1)$ and $Y \sim Unif[0,2]$, then

$$f_X(x) = \begin{cases} e^{-x} & \text{if } x \geq 0 \\ 0 & \text{otherwise}, \end{cases} \quad \text{and} \quad f_Y(y) = \begin{cases} \frac{1}{2} & \text{if } 0 \leq y \leq 2 \\ 0 & \text{otherwise}. \end{cases}$$

If X and Y are also independent, then

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) = \begin{cases} \frac{1}{2}e^{-x} & \text{if } x \geq 0,\ 0 \leq y \leq 2 \\ 0 & \text{otherwise}. \end{cases}$$

- Covariance, variance, expectation, correlation (rules & definitions)

$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ always

$Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y)$ always

$Var(X+Y) = Var(X) + Var(Y)$ if X & Y are independent (or uncorrelated)

Def: $Cov(X,Y) = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]$
$= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$

Def: $\text{Correlation}(X,Y) = \dfrac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$

4. Limit theorems

If $X_1, X_2, \ldots$ are iid with $\mathbb{E}[X_i] = m$, $Var(X_i) = \sigma^2 > 0$, then for n large:

LLN: $\frac{1}{n}\sum_{i=1}^{n} X_i \approx m$

CLT: $\frac{1}{n}\sum_{i=1}^{n} X_i \approx m + \frac{\sigma}{\sqrt{n}} Z$ where $Z \sim N(0,1)$,

standard normal

or equivalently $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{X_i - m}{\sigma}\right) \approx Z$.

Ex: $\mathbb{P}\left(a \leq \frac{1}{n}\sum_{i=1}^{n} X_i \leq b\right) \approx \mathbb{P}\left(a \leq m + \frac{\sigma}{\sqrt{n}} Z \leq b\right) = \ldots$