Central limit theorem

Last time inequalities

Markov: $X \geq 0$ then $P\{X \geq a\} \leq \frac{E[X]}{a}$

Chebyshev: $P\{|X-\mu| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}$

Weak law of large numbers

$X_1, X_2, \dots,$ independent identically distributed

then
$$\lim_{n\to\infty} P\left\{\left|\frac{X_1+\dots+X_n}{n} - \mu\right| \geq \varepsilon\right\} = 0$$

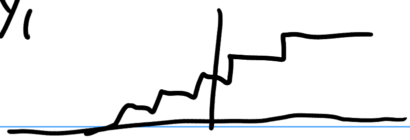
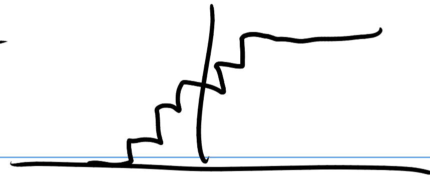
The weak law of large numbers (and the Central limit theorem) are statements about convergence in distribution:

Let Y_n be a sequence of Random variables with CDF's $F_{Y_n}(a)$

If Y is a random variable with $F_Y(a)$

We say the sequence Y_n converges to Y in distribution if

$\lim_{n\to\infty} F_{Y_n}(a) = F_Y(a)$ For every a at which F_Y is continuous

F_{Y_1}  , F_{Y_2}  , F_{Y_3} 

F_Y 

Apply this to weak law of large numbers

$$Y_n = \frac{X_1 + \cdots + X_n}{n} \qquad E[Y_n] = \mu$$

$$Var(Y_n) = \frac{\sigma^2}{n}$$

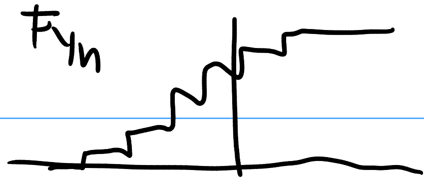
$$\lim_{n \to \infty} P\left\{ \frac{X_1 + \cdots + X_n}{n} \leq \mu - \varepsilon \right\} = 0$$

$$\lim_{n \to \infty} P\{Y_n \leq \mu - \varepsilon\} = 0 \qquad \text{for any } \varepsilon > 0$$

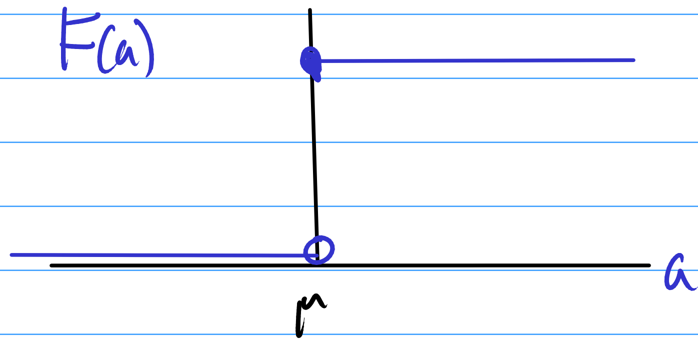
similarly $\displaystyle\lim_{n \to \infty} P\left\{ \frac{X_1 + \cdots + X_n}{n} \leq \mu + \varepsilon \right\} = 1$

$$\lim_{n \to \infty} P\{Y_n \leq \mu + \varepsilon\} = 1$$

$$\lim_{n \to \infty} F_{Y_n}(\mu - \varepsilon) = 0 \qquad \lim_{n \to \infty} F_{Y_n}(\mu + \varepsilon) = 1$$



limiting distribution $F(a) = \begin{cases} 0 & \text{if } a<\mu \\ 1 & \text{if } a \geq \mu \end{cases}$



The step function is the CDF of a constant random variable.

$Y=\mu$ with probability 1

then $F_Y(a) = \begin{cases} 0 & a<\mu \\ 1 & a \geq \mu \end{cases}$

$$Y_n = \frac{X_1 + \cdots + X_n}{n} \xrightarrow{\text{distribution}} \mu \text{ (constant)}$$

The central limit theorem says that Y_n is asymptotic to a sequence of normal random variables.

$$Y_n = \frac{X_1 + \cdots + X_n}{n} \qquad E[Y_n] = \mu \qquad Var(Y_n) = \frac{\sigma^2}{n}$$

$$Z_n = \frac{\left(\frac{X_1 + \cdots + X_n}{n} - \mu\right)}{\left(\frac{\sigma}{\sqrt{n}}\right)} = \frac{(X_1 + \cdots + X_n - n\mu)}{\sigma\sqrt{n}}$$

Z_n is like the average, but it has mean 0 and variance instead.

Let Z be standard normal Random variable

$$F_Z(a) = \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-t^2/2}\, dt$$

Theorem (CLT): $Z_n \xrightarrow{\text{distribution}} Z$.

i.e., $$\lim_{n\to\infty} P\left\{ a \le \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \le b \right\} = \Phi(b) - \Phi(a)$$

* Remember, this is all under the assumption
* that $X_1, X_2, \ldots$ are independent and
* have a common distribution w/ finite μ, σ^2

Other ways to phrase it

For n large

$$X_1 + \cdots + X_n \overset{\text{dist.}}{\sim} \text{normal with mean } n\mu \text{ and variance } n\sigma^2$$

$$\frac{X_1 + \cdots + X_n}{n} \overset{\text{dist}}{\sim} \text{normal with mean } \mu \text{ and variance } \sigma^2/n$$

Extension to when variables are not identically distributed, but independent.

$$X_1 + \cdots + X_n \quad \text{mean } \sum_{i=1}^{n} \mu_i$$

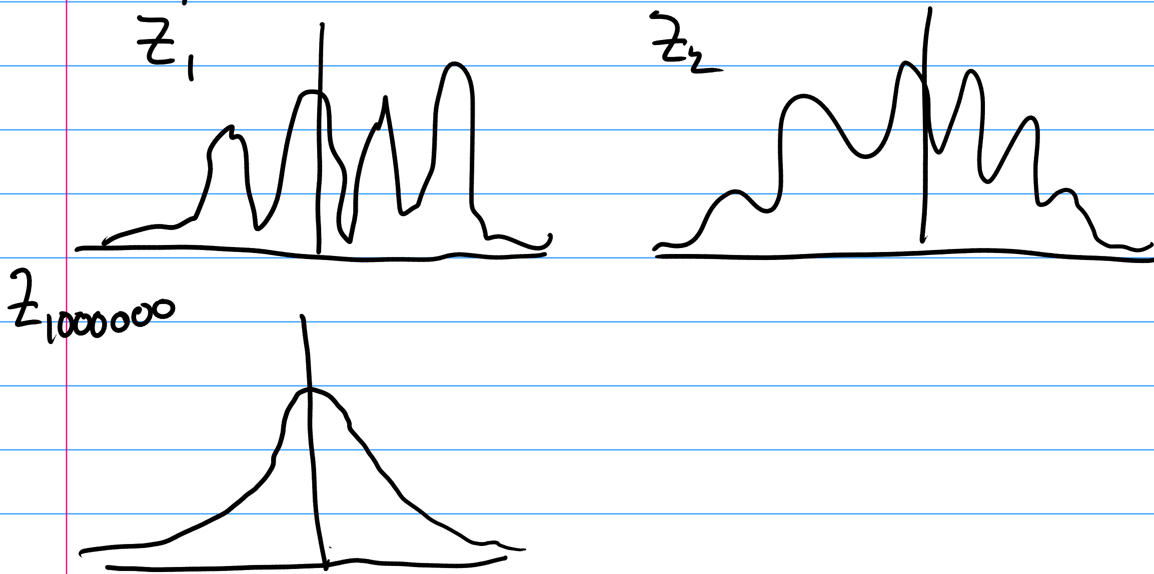
$$\text{variance } \sum_{i=1}^{n} \sigma_i^2$$

$$\frac{\sum_{i=1}^{n} (X_i - \mu_i)}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}} \overset{\text{dist}}{\sim} \text{normal with mean } 0 \text{ and variance } 1$$

(Under some technical hypotheses)

This theorem explains the ubiquity of the normal distribution: Appears whenever we look at a Random variable which is a sum of many small independent contributions.

In pictures of PDF



Application of CLT: compute approximate probabilities.

X is binomial with (n, p)

$$X = \sum_{i=1}^{n} X_i$$

X_i is binomial with $(n=1, p)$

$$E[X_i] = p$$
$$Var(X_i) = p(1-p)$$

Large $n \Rightarrow$

$$\frac{X - np}{\sqrt{np(1-p)}}$$ is approximately standard normal

This is just DeMoivre-Laplace limit theorem

(for this approximation would use the continuity correction!)

X = Poisson with parameter $\lambda = 100$

$$X = \sum_{i=1}^{100} X_i \qquad X_i \text{ is Poisson with parameter } \lambda = 1$$

Mean $(X) = \lambda = 100$ Mean $(X_i) = \lambda = 1$

Var $(X) = \lambda = 100$ Var $(X_i) = \lambda = 1$

So we can approximate X by a normal distribution with $\mu = 100$ $\sigma^2 = 100$

Q: $$P\{X \geq 120\} = e^{-100} \sum_{i=120}^{\infty} \frac{(100)^i}{i!}$$

Because we're approximating a discrete distribution by a continuous one, we use the continuity correction

$$P\{X \geq 120\} = P\{X > 119.5\}$$

$$\overset{CLT}{\approx} P\left\{Z > \frac{119.5 - 100}{\sqrt{100}}\right\} = P\{Z > 1.95\}$$

$$= 1 - \Phi(1.95) = .0256$$

If X is an integer (continuity correction)

$$\begin{cases} X \leq n \iff X < n + \frac{1}{2} \\ X < n \iff X < n - \frac{1}{2} \\ X \geq n \iff X > n - \frac{1}{2} \\ X > n \iff X > n + \frac{1}{2} \end{cases}$$