Probability and Statistics

4 – Continuous Random Variables

Stefan Heiss

Technische Hochschule Ostwestfalen-Lippe
Dep. of Electrical Engineering and Computer Science

December 15, 2023

PDF's for Sums of Random Variables

Theorem (4.32)

If X and Y are independent random variables with pdf's f_X and f_Y, respectively, then the convolution of f_X and f_Y is a pdf of $X+Y$:

$$f_{X+Y}(z) = (f_X * f_Y)(z) = \int_{-\infty}^{\infty} f_X(x) \cdot f_Y(z-x)\, dx$$

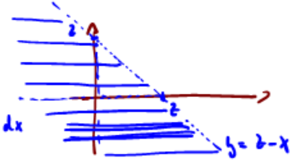
Pf: $Z = X+Y$

cdf of (X,Y): $f_{XY}(x,y) = f_X(x) \cdot f_Y(y)$

because X and Y are independent

$F_Z(z) = Pr(X+Y \leq z)$

$y \leq z - x$



$$f_Z(z) = \frac{d}{dz}\int_{-\infty}^{\infty}\int_{-\infty}^{z-x} f_X(x)\cdot f_Y(y)\, dy\, dx = \int_{-\infty}^{\infty}\left(\frac{d}{dz}\int_{-\infty}^{z-x} f_X(x)\, f_Y(y)\, dy\right) dx = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\, dx \quad \square$$

Central Limit Theorem

$X_1, X_2, \ldots$ $E(X_i) = \mu$, $Var(X_i) = \sigma^2$

Theorem (4.54)

Let $(X_i)_{i \in \mathbb{N}}$ be a sequence of independent, identically distributed random variables with finite mean μ and finite variance σ^2. Let F_{Y_n} be the cdf of

$$Y_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right) = \frac{\left(\sum_{i=1}^{n} X_i\right)/n - \mu}{\sigma/\sqrt{n}}$$

Then $E(Y_n) = 0$, $Var(Y_n) = 1$ for all $n \in \mathbb{N}$ and:

$$\lim_{n \to \infty} F_{Y_n} = \Phi$$

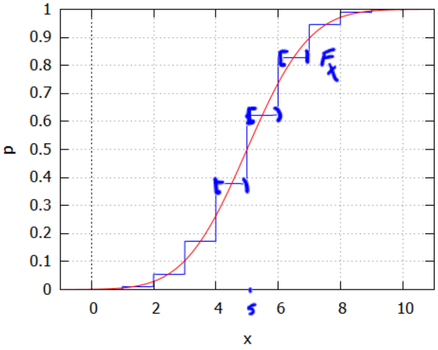
Central Limit Theorem - Example: $X_i \sim \mathrm{Bernoulli}(p)$

- $\displaystyle\sum_{i=1}^{n} X_i \sim \mathrm{binomial}(n, p)$
- $\displaystyle Y_n = \frac{\left(\sum_{i=1}^{n} X_i\right)/n \;-\; \mu}{\sigma/\sqrt{n}} = \frac{\left(\sum_{i=1}^{n} X_i\right)/n \;-\; p}{\sqrt{p(1-p)}/\sqrt{n}} = \frac{\left(\sum_{i=1}^{n} X_i\right) \;-\; np}{\sqrt{np(1-p)}} \;\overset{\sim}{\to}\; \mathcal{N}(0,1)$
- $\displaystyle\sum_{i=1}^{n} X_i \;\overset{\sim}{\to}\; \mathcal{N}\left(np, \sqrt{np(1-p)}\right)$

Central Limit Theorem - Example: $X_i \sim \mathrm{Bernoulli}(p)$

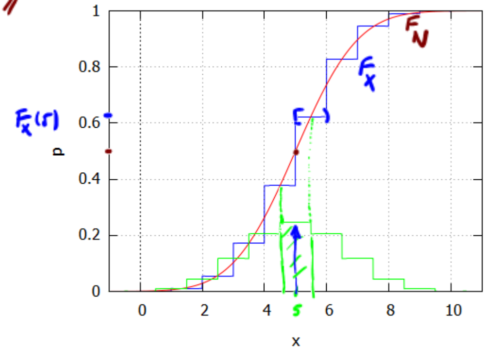
- $X = \sum_{i=1}^{n} X_i \sim \mathrm{binomial}(n, p) \quad \rightarrow \quad \mathcal{N}(np, \sqrt{np(1-p)})$

n=10
p=0.5



Continuity Correction

- $\sum_{i=1}^{n} X_i \sim \text{binomial}(n, p) \;\to\; \mathcal{N}(np, \sqrt{np(1-p)})$



Eg. $Pr(X \leq 5) = F_X(5)$

$F_N(5)$

better: $\underline{F_N(5.5)}$

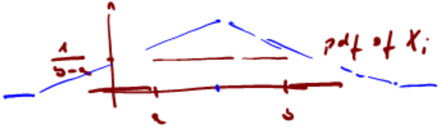
Central Limit Theorem - Example: $X_i \sim \mathrm{geometric}(p)$

- $\displaystyle\sum_{i=1}^{n} X_i \sim \mathrm{nbino}(n, p)$
- $Y_n = \dfrac{\left(\sum_{i=1}^{n} X_i\right)/n - \mu}{\sigma/\sqrt{n}} = \dfrac{\left(\sum_{i=1}^{n} X_i\right)/n - 1/p}{(\sqrt{1-p}/p)/\sqrt{n}} = \dfrac{\left(\sum_{i=1}^{n} X_i\right) - n/p}{\sqrt{n(1-p)}/p} \overset{\sim}{\to} \mathcal{N}(0,1)$
- $\displaystyle\sum_{i=1}^{n} X_i \overset{\sim}{\to} \mathcal{N}\left(n/p, \sqrt{n(1-p)}/p\right)$

Central Limit Theorem - Example: $X_i \sim \mathrm{Poisson}(\lambda)$

- $\displaystyle\sum_{i=1}^{n} X_i \sim \mathrm{Poisson}(n\lambda)$
- $Y_n = \dfrac{\left(\sum_{i=1}^{n} X_i\right)/n - \mu}{\sigma/\sqrt{n}} = \dfrac{\left(\sum_{i=1}^{n} X_i\right)/n - \lambda}{\sqrt{\lambda}/\sqrt{n}} = \dfrac{\left(\sum_{i=1}^{n} X_i\right) - n\lambda}{\sqrt{n\lambda}} \overset{\sim}{\to} \mathcal{N}(0,1)$
- $\displaystyle\sum_{i=1}^{n} X_i \overset{\sim}{\to} \mathcal{N}\left(n\lambda, \sqrt{n\lambda}\right)$

Central Limit Theorem - Example: $X_i \sim \mathrm{uniform}[a, b]$



- $\sum_{i=1}^{n} X_i \sim$???

$X_1 + X_2$

$(X_1 + X_2) + (X_3 + X_4)$

$\vdots$

See Matlab script

Central Limit Theorem - Example: $X_i \sim \mathrm{uniform}[a, b]$

- $$\sum_{i=1}^{n} X_i \sim ???$$
- $$Y_n = \frac{\left(\sum_{i=1}^{n} X_i\right)/n - \mu}{\sigma/\sqrt{n}} = \frac{\left(\sum_{i=1}^{n} X_i\right)/n - \frac{a+b}{2}}{\frac{b-a}{\sqrt{12}} \cdot \frac{1}{\sqrt{n}}} = \frac{\left(\sum_{i=1}^{n} X_i\right) - \frac{n(a+b)}{2}}{\sqrt{n} \cdot \frac{b-a}{\sqrt{12}}} \overset{\sim}{\to} \mathcal{N}(0,1)$$
- $$\sum_{i=1}^{n} X_i \overset{\sim}{\to} \mathcal{N}\left(\frac{n(a+b)}{2}, \sqrt{n} \cdot \frac{b-a}{\sqrt{12}}\right)$$

Central Limit Theorem - Example: $X_i \sim \exp(\lambda)$

- $\displaystyle\sum_{i=1}^{n} X_i \sim ???$ (Erlang distribution)

- $Y_n = \dfrac{(\sum_{i=1}^n X_i)/n - \mu}{\sigma/\sqrt{n}} = \dfrac{(\sum_{i=1}^n X_i)/n - 1/\lambda}{(1/\lambda)/\sqrt{n}} = \dfrac{(\sum_{i=1}^n X_i) - n/\lambda}{\sqrt{n}/\lambda} \overset{\sim}{\to} \mathcal{N}(0,1)$

- $\displaystyle\sum_{i=1}^{n} X_i \overset{\sim}{\to} \mathcal{N}\left(n/\lambda, \sqrt{n}/\lambda\right)$

Central Limit Theorem - Example: $X_i \sim \mathcal{N}(\mu, \sigma)$

- $\displaystyle\sum_{i=1}^{n} X_i \sim \mathcal{N}(n\mu, \sqrt{n}\,\sigma)$
- $Y_n = \dfrac{\left(\sum_{i=1}^{n} X_i\right)/n \; - \; \mu}{\sigma/\sqrt{n}} = \dfrac{\left(\sum_{i=1}^{n} X_i\right) \; - \; n\mu}{\sqrt{n}\,\sigma} \sim \mathcal{N}(0, 1)$