$\underline{X}$ random variable

↳ discrete → range is either a finite or countable set

continuous → range is every value in some interval $[a,b]$

discrete random variable

$\{x_1, \dots, \}$

$p(x_i) = P(\underline{X} = x_i)$ (p.m.f)

continuous random variable

$$\int_a^b f(x)\,dx = P(\underline{X} \in (a,b))$$

→ cdf $F(x) = P(\underline{X} \leq x)$

$$\Rightarrow \sum_{x_i \leq x} p(x_i)$$

$$\Rightarrow \int_{-\infty}^{x} f(y)\,dy$$

in the continuous, what is the relationship between the cdf and pdf

$\Rightarrow f = \frac{d}{dx} F$

Suppose we have two random variables that are both defined on the same sample space, X and Y $\Rightarrow$ their values are both functions of outcomes of the same experiment.

both discrete

joint pmf $p(x, y) = p(X = x \cap Y = y)$

Ex Toss two coins

$X = \begin{cases} 1, & H \\ 0, & T \end{cases} \qquad Z = \begin{cases} 1, & H \\ 0, & T \end{cases}$

$Y = X + Z$

$\Rightarrow$ if I have information about X, I have information about the possible values of Y.

$\Rightarrow$ if $X = 1$, then Y can't be zero

encode all of this information in
the joint pmf

$\Rightarrow$ two continuous random variables

$f(x,y)$

$P(X \in [a,b], Y \in [c,d])$

$$= \int_a^b \int_c^d f(x,y)\, dy\, dx$$

diff $x \in [a,b]$ and $x \in (a,b)$

$\hookrightarrow \{x \in (a,b)\} \cup \underline{\{x = a\} \cup \{x = b\}}$

$\Rightarrow$

Independent Random

$X, Y \Rightarrow$ independent it means that
information about X doesn't give
us any info about Y

$$p(x,y) = p_X(x) \cdot p_Y(y)$$

$$f(x,y) = f_X(x) \cdot f_Y(y)$$

Expectation

↳ expectation of a random variable X can be thought of prob-weighted average value

$X \to \{x_1, \ldots,\}$ (discrete)

$$E[X] = \sum_{x_i} p(x_i) \cdot x_i$$

$$= \int_{-\infty}^{\infty} x f(x) \, dx$$

Var(X) → the probability-weighted average squared distance from the expectation

$\mu = E[X]$

$$E[(X-\mu)^2]$$

$$= E[X^2] - E[X]^2$$

Covariance (X, Y)

↳ the probability average product of X's signed distance from its

expectation with Y's signed distance
from it's expectation

→ if $X > \mu_X$ at the same $Y > \mu_Y$
most of the time
$Cov(X, Y)$ will be positive

<u>Units of variance</u> → $(\text{units of } X)^2$

$$SD(X) = \sqrt{Var(X)}$$

↳ converts back to the units of
X

<u>Units of $Cov(X, Y)$</u> → units of x × units of
y

$$Corr(X, Y) = \frac{Cov(X, Y)}{SD(X) \cdot SD(Y)}$$

<u>Properties</u>

① $E[aX+b] = aE[X] + E[b]$

$= aE[X] + b$

② $E[g(X)] = \sum g(x) \cdot p(x)$

$$= \int g(x) \cdot f(x)\,dx$$

③ $$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) \cdot f(x,y)\,dx\,dy$$

$$Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$$

$$\mu_x = E[X]$$

$$\mu_y = E[Y]$$

$$= \iint (x-\mu_x)(y-\mu_y) f(x,y)\,dy\,dx$$

④ $$\underline{E[X+Y]} = \iint (x+y) f(x,y)\,dx\,dy$$

$f(x,y)$ $$= \iint x \cdot f(x,y)\,dx\,dy$$

$$+ \iint y \cdot f(x,y)\,dx\,dy$$

$$= E[X] + E[Y]$$

⑤ $$Var(aX+b) = a^2 Var(X)$$

⑥ $\mathrm{Cov}(X, Y) = 0$

X & Y are independent

⑦ $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X, Y)$

Week Law of Large Numbers

$X_1, \ldots, X_n$ is a sequence of independent identically distributed random variables.

$E[X_i] = \mu$ and $\mathrm{Var}(X_i) = \sigma^2$

$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$.

$\varepsilon > 0$

$\lim_{n \to \infty} P(|\overline{X}_n - \mu| > \varepsilon) \to 0$

$\mu = 60$ $\quad \varepsilon > .01$

$n =$ a billion

$P(\overline{X}_{1\text{billion}} \notin (59.99, 60.01))$

≈ 0.0

Families of distribution

remember that pmf $p(\cdot)$

pdf $f(\cdot)$

related group of functions that can be indexed by a parameter that each correspond to a type of probability

① Bernoulli random variable:

$X \to \{0,1\}$, parameter p

$p(0) = 1-p$ $\quad p \in (0,1)$

$p(1) = p$

② Binomial random variable parameters (n,p)

X = sum of n independent Bernoulli (p)

range of $X \to \{0, \dots, n\}$

$$p(i) = \binom{n}{i} p^i (1-p)^{n-i}$$

$$\binom{n}{i} = \frac{n!}{i!(n-i)!}$$

} positive but bounded integers

$\underline{n, p}$ $\quad n \in \mathbb{Z}^{+}$

$p \in (0, 1)$

↳ species a pmf.

(3) Poisson random variables $(\lambda > 0)$

$X \sim \text{Poiss}(\lambda)$

$X \to \{0, 1, 2, \dots\}$

$$p(i) = \frac{e^{-\lambda} \lambda^{i}}{i!}$$

↳ positive but unbounded integers.

$E[X] = \lambda$

$\text{Var}[X] = \lambda$

(4) Uniform distribution on $[\underline{a}, \underline{b}]$

↳ continuous random variable with range $[a, b]$.

↳ we don't believe any value in the range is more likely than another

$$f_X(x) = \begin{cases} 1/_{b-a}, & x \in [a, b] \\ 0, & \text{otherwise} \end{cases}$$

⑤ Normal distribution (μ, σ)

$$X \to (-\infty, \infty)$$

$$f_X(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$E[X] = \mu$$

$$Var(X) = \sigma^2$$

⑥ Exponential distribution

$$X \to (0, \infty) \quad (\lambda > 0)$$

$$f_X(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x > 0 \end{cases}$$

$$E[X] = \frac{1}{\lambda} \quad Var(X) = \frac{1}{\lambda^2}$$

Central Limit Theorem

$X_1, \ldots, X_n$ iid random variables

with $E[X_i] = \mu$ and $Var(X_i) = \sigma^2$

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

$$E[\bar{X}_n] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right]$$

$$= \frac{1}{n}\sum E[X_i]$$

$$= \frac{1}{n}\cdot n\mu = \mu$$

$$Var(\bar{X}_n) = \frac{\sigma^2}{n}$$

$$\underline{Z_n = \sqrt{n}\cdot\frac{\bar{X}_n - \mu}{\sigma}} \rightarrow N(0,1)$$

$\Rightarrow$ if n large (?)

$$\bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

given $\varepsilon > 0$

$$P(|\bar{X}_n - \mu| > \varepsilon) \approx P$$

n large is probably about 30