Lecture

For probability space: (Ω, F, P), r.v $X : \Omega \to \mathbb{R}$

- Distribution of X (= "frequencies w which X takes values x") captured by its cdf:

$$F_X(x) := P\{X \le x\} \qquad x \in \mathbb{R}$$

- discrete rv's : F_X is a "step fcn" w jumps

$$P_X(x) = F_X(x) - F_X(x^-)$$

- cts rv's : F_X of form

$$F_X(x) = \int_{-\infty}^{\infty} \underbrace{f_X(u)\,du}_{\text{pdf for } X} \qquad \begin{pmatrix} f_X \geq 0 \\ \int f_X = 1 \end{pmatrix}$$

- computing expectations, etc are basically the same in both cases

  ↳ discrete : pdf & sums

  ↳ cts : pdf & integraal

- 3 examples of cts. distributions: Unif(a, b), Exp(N), $\mathcal{N}(\mu, \sigma^2)$

Origin of Exp(λ):

- Suppose we want to model a memoryless process
- Mathematically: Let X be a non-negative rv with a memoryless property:

$$P\{X > t + s \mid X > s\} = P\{X > t\} \qquad \forall t, s \geq 0$$

$$P(\{X > t + s\} \cap \{\cancel{X > s}\}) = P\{X > t\}\, P\{X > s\}$$

$$\hookrightarrow = P\{X > t + s\}$$

complementary CDF $\bar{F}_X = 1 - F_X$

$$\bar{F}_X(t+s) = \bar{F}_X(t)\,\bar{F}_X(s)$$ the unique soln to this that is a cdf of the form $\bar{F}_X(x) = e^{-\lambda x}, \ \lambda > 0$

- If X has memoryless property,

$$F_X(x) = 1 - e^{-\lambda t} \quad \text{for some } \lambda > 0$$

$$\Rightarrow X \sim Exp(\lambda)$$

Gaussian Random Variables

- In a sense to be made precise later, the Gaussian distribution emerges as the "sum of independent effects"
  - ↳ think pollen placed into water
- We call $\mathcal{N}(0,1)$ the standard normal distribution.

cdf : $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\{-u^2/2\} du$

ex:

$\mathcal{N}(\mu, \sigma^2)$

$$F_X(x) = P\{X \leq x\} = P\Big(\underbrace{\frac{X-\mu}{\sigma}}_{\sim \mathcal{N}(0,1)} \leq \frac{x-\mu}{\sigma}\Big) = \Phi\Big(\frac{x-\mu}{\sigma}\Big)$$

- Gaussians have many nice properties
  - ↳ If X gaussian, then so is $aX + b$
  - ↳ If X, Y are indep. gaussians, then so is $\underbrace{X + Y}_{\text{params: } \mu_X + \mu_Y,\ Var(X) + Var(Y)}$

Ex: Let $V \sim \mathcal{N}(1, \sigma^2)$ be input voltage to some chip avg'd over 1 sec.

↳ Chip fails if voltage dips below 0.5 V or exceeds 2.5 V for a 1-sec period:

$P(\text{chip fails in 60s duration})$

$$\leq 60 P(\text{chip fails in 1sec})$$

$$\leq 60(\underbrace{P(V \leq 0.5)} + \underbrace{P(V \geq 2.5)})$$

$$= 60\Big(\Phi\Big(\frac{0.5-1}{\sigma}\Big) + \Big(1 - \Phi\Big(\frac{2.5-1}{\sigma}\Big)\Big)\Big)$$

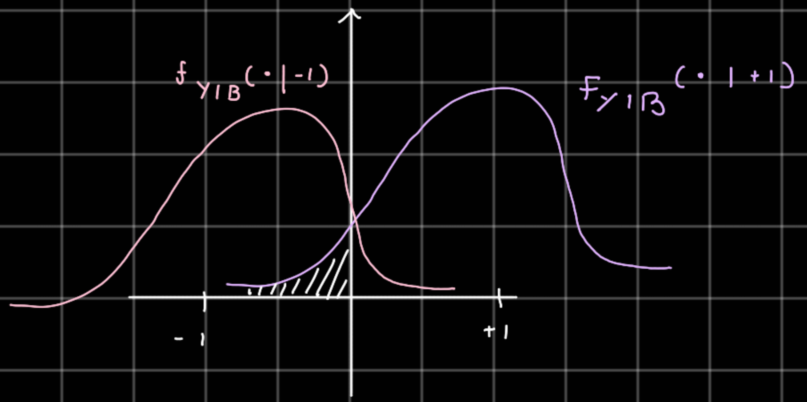
Ex: cellphone sends a bit $B \in \{\pm 1\}$ tower

Tower receives $Y = B + N$, $N \sim \mathcal{N}(0,1)$

and makes decision

$$\hat{B}(Y) = \text{sign}(Y)$$

$$P(\text{error} \mid B = +1) = P(Y > 0 \mid B = +1) = \Phi(-1)$$



Q/ Given I receive $\{Y = y\}$, what is probability that $B = +1$?

↳ problem: Y is cts, so $P(Y=y) = 0$

$$P(\{B=+1\} \cap \{Y \in [y, y+\delta]\}) = \underbrace{P(Y \in [y, y+\delta] \mid B=+1)}_{\sim f_{Y|B}(y|+1)\,\delta} P_B(+1)$$

$$= P(B=+1 \mid Y \in [y, y+\delta])\underbrace{P(Y \in [y, y+\delta])}_{\sim f_Y(y)\,\delta}$$

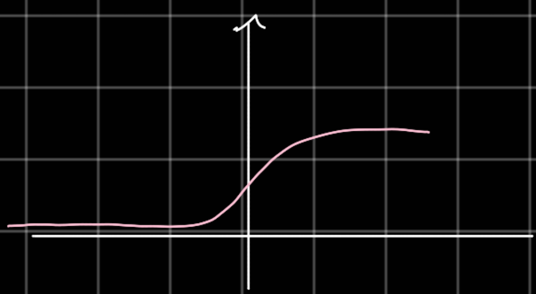
$$P(B=+1 \mid Y \in [y, y+\delta])\, f_Y(y)\cancel{\delta} = f_{Y|B}(y|+1)\cancel{\delta}\, P_B(+1)$$

$$\Rightarrow P_{B|Y}(B=+1 \mid y) = \frac{P_B(+1)\, f_{Y|B}(y|+1)}{f_Y(y)}$$

Assume $P_B(b) = \frac{1}{2}$, $b = \pm 1$

$$P_{B|Y}(+1|y) = \frac{\cancel{\frac{1}{2}}\cancel{\frac{1}{\sqrt{2\pi}}}\exp\left(-\frac{(y-1)^2}{2}\right)}{\cancel{\frac{1}{2}}\cancel{\frac{1}{\sqrt{2\pi}}}\exp\left(-\frac{(y+1)^2}{2}\right) + \cancel{\frac{1}{2}}\cancel{\frac{1}{\sqrt{2\pi}}}\exp\left(-\frac{(y-1)^2}{2}\right)}$$

$$= \frac{1}{1 + e^{-2y}} \rightarrow$$



Last example motivates the definition of cond. density as follows:

If X, Y jointly continuous, define conditional density of X given Y by

$$f_{X|Y}(x,y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

Interpretation: $f_{X|Y}(\cdot \mid y)$ is the density of X given I know $\{Y=y\}$

ex: Bayes rule

$$f_{X|Y}(x|y) = \frac{f_X(x)}{f_Y(y)} f_{Y|X}(y|x)$$

Ex: For X, Y jointly continuous, we have:

$$\mathbb{E}[X \mid Y=y] = \int x \, f_{X|Y}(x|y)\, dx$$

↳ conditional expectation still satisfies tower property:

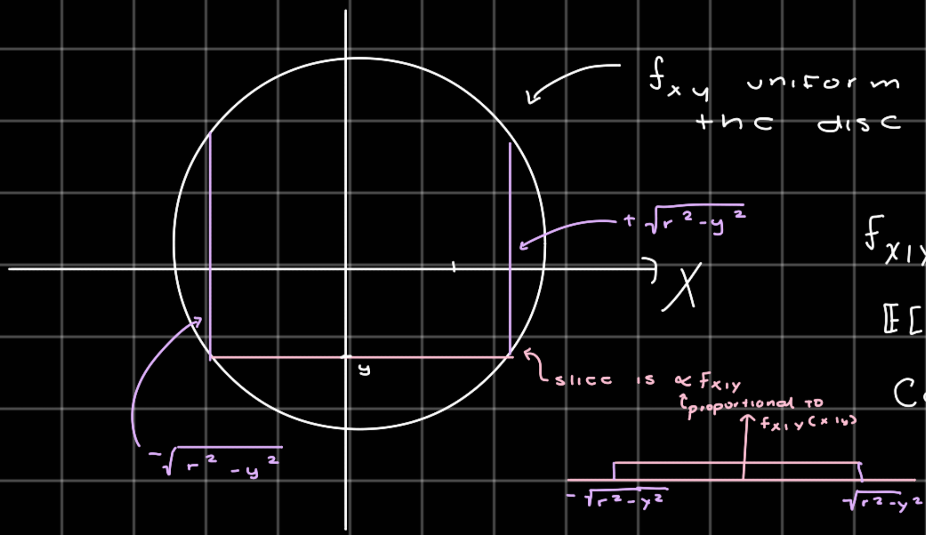
$$\mathbb{E}[g(Y)X] = \mathbb{E}[g(Y)\,\mathbb{E}[X \mid Y]]$$

↑ $\mathbb{E}[X|Y=y]$ evaluated at Y

ex: $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]$

- Oftentimes for calculations, helps to draw a picture ("graphical density")

Ex: dartboard example



$f_{X|Y}(x|y)$

$\mathbb{E}[X|Y=y] = 0$

$$\begin{aligned} \mathrm{Cov}(X,Y) &= \mathbb{E}[XY] \\ &= \mathbb{E}[\mathbb{E}[X|Y]Y] \\ &= 0 \end{aligned}$$

Derived Distributions

- distribution of X is given by cdf F_X. Suppose we define $Y = g(X)$ for some f'n g.

Q/ What's the distribution of Y?

① Ask yourself: Do you really need this distr of Y?

↳ ex, can use LOTUS for

$$\mathbb{E}[f(Y)] = \mathbb{E}[f(g(X))]$$ can be computed w/ distr. of X

② If you really need it, best to work w/ CDFs.

$$F_Y(y) = P\{Y \le y\}$$
$$= P\{g(X) \le y\}$$
$$= P\{X \in g^{-1}((-\infty, y))\}$$

inverse image of g $\{x : g(x) \le y\}$

No closed form in general

Ex:

X cts rv $Y = aX + b$, $a \neq 0$

$$F_Y(y) = P\{aX \le y - b\}$$
$$= \begin{cases} P\{X \le \frac{y-b}{a}\} & a > 0 \\ P\{X \ge \frac{y-b}{a}\} & a < 0 \end{cases}$$
$$= \begin{cases} F_X\left(\frac{y-b}{a}\right) & a > 0 \\ 1 - F_X\left(\frac{y-b}{a}\right) & a < 0 \end{cases}$$

$$f_Y(y) = F_Y'(y) = \begin{cases} \frac{1}{a} F_X'\left(\frac{y-b}{a}\right) & a > 0 \\ -\frac{1}{a} F_X'\left(\frac{y-b}{a}\right) & a < 0 \end{cases}$$
$$= \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

↳ in general:

$Y = AX + \vec{b}$, A invertible

$$f_Y(y) = \frac{1}{|\det(A)|} f_X\left(A^{-1}(y - b)\right)$$