A **stochastic process** is a collection of random variables

$\{X_t\}_{t \in \{0,1,2,3,\dots\}}$ or $\{X_t\}_{t \geq 0}$

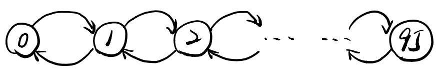
indexed by time t.

In this class, X_t will be discrete.

e.g. <u>class enrollment</u>

time 0 registration opens

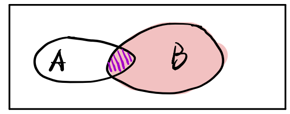
students begin to add or drop $X_t \in \{0, 1, \dots, 95\}$, $t \in [0, 200]$



Review

restricting sample space

Conditional probability



$P(A|B) = P(A\cap B)/P(B)$

$P(A\cap B) = P(A|B)P(B) = P(B)P(A|B)$

Law of Total Probability

$P(A) = P(A\cap B) + P(A\cap B^c)$ — mutually exclusive; $B + B^c$ entire sample space

$= P(A|B)P(B) + P(A|B^c)P(B^c)$

e.g. Patients recruited to two arms (control, treatment)

Probability of an adverse event on control arm is P_C.
Probability of an adverse event on the treatment is P_T.

What the probability that adverse event occurs first on control arms?

Define F: event that adverse event occurs first on the control arm

E_1: adverse event occurs on the first control patient

E_2: adverse event doesn't occur on the first control patient but occur on first treatment patient

E_3: neither first patient has the adverse event.

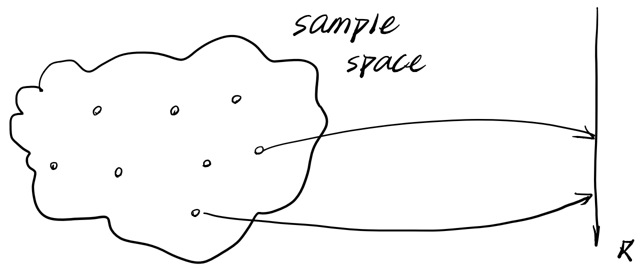
$$\Rightarrow P(F) = \underbrace{P(F|E_1)}_{1}\underbrace{P(E_1)}_{P_C} + \underbrace{P(F|E_2)}_{0}\underbrace{P(E_2)}_{P_T(1-P_C)} + \overbrace{P(F|E_3)}^{P(F)}\underbrace{P(E_3)}_{(1-P_C)(1-P_T)}$$

$$= P_C + P(F)(1-P_C)(1-P_T)$$

$$\Rightarrow P(F) = \frac{P_C}{P_C + P_T(1-P_C)}$$

Random variable

A map from sample space to a real line.



① Discrete random variables

Assign a positive to a countable number of values (possibly infinite)

Probability mass function

$$P(X_{t+1}=x_{t+1} \mid X_t = x_t) = \frac{P(X_t = x_t, X_{t+1} = x_{t+1})}{P(X_t = x_t)}$$

$\sum_s P(X_{t+1}=s)\, P(X_t = x_t \mid X_{t+1}=s)$

② Continuous random variables

Allow for a uncountable number of possible values.

Probability density function (define probability on intervals)

$$f_{X_{t+1}, X_t}(x_{t+1} \mid x_t) = \frac{f_{X_{t+1}, X_t}(x_t, x_{t+1})}{f_{X_t}(x_t)}$$

Expectation

$$E(Y) = \begin{cases} \sum y\, P(Y=y) \\ \int_{-\infty}^{\infty} y\, f_Y(y)\, dy \end{cases}$$

$X = g(Y)$

$$\Rightarrow E(X) = \sum x\, P(X=x)$$
$$= \sum x\, P(g(Y)=x)$$
$$= \sum x\, P(y \in \{y : y = g^{-1}(x)\})$$

Conditional Expectation

$$E(Y|X=x) = \begin{cases} \sum_y y\, P(Y=y|X=x) \\ \int_{-\infty}^{\infty} y\, f_{Y|X}(y|x)\, dy \end{cases}$$

a function of x (little x)
unless X, Y are independent

① $E(aY+b \mid X=x) = aE(Y \mid X=x)+b$

② $E(g(Y) \mid X=x) = \begin{cases} \sum g(y)\, P(Y=y \mid X=x) \\ \int_{-\infty}^{\infty} g(y) f_{Y|X}(y|x)\, dy \end{cases}$

③ $E(Y|X=x) = E(Y)\ \forall x$ iff Y and X are independent

④ If $Y=g(X)$, $E(Y \mid X=x) = g(x)$

Law of Iterated Expectation

$$E_Y(Y) = E_X[E_{Y|X}[Y|X]]$$

not a function of Y
but a function of X (big X)
$\Rightarrow$ a random variable

$$RHS = \sum_x E_{Y|X}[Y|X=x]\,P(X=x) = \sum_x \Big[\sum_y y\,P(Y=y|X=x)\Big] P(X=x)$$
$$= \sum_y \Big[\sum_x y\,P(Y=y|X=x)\Big] P(X=x) = \sum_y y\, \underbrace{\Big[\sum_x P(Y=y|X=x)\,P(X=x)\Big]}_{=P(Y=y)}$$
$$= \sum_y y\,P(Y=y) = E_Y(Y) = LHS$$

e.g. $T = \sum_{i=1}^{N} X_i$ where $X_i \overset{iid}{\sim} Poisson(\lambda)$
Find $E_T(T)$ if N is random.

$$E_{T|N=n}(T \mid N=n) = E_{T|N=n} \sum_{i=1}^{n} X_i = n\lambda$$
$$E_T(T) = E_N[E_{T|N}(T|N)] = E_N[N\lambda] = \lambda E_N[N] = \lambda \mu_N$$

Law of Iterated Variance

$$Var_Y(Y) = \underbrace{Var_X(E_{Y|X}[Y|X])}_{①} + \underbrace{E_X(Var_{Y|X}[Y|X])}_{②}$$

Note: $Var(W) = E_W(W^2) - [E_W(W)]^2$
$= E_W[(W - E_W(W))^2]$

① $Var_X(E_{Y|X}[Y|X]) = E_X[E_{Y|X}[Y|X]^2] - \underbrace{(E_X[E_{Y|X}[Y|X]])^2}_{=[E_Y(Y)]^2}$

② $E_X(Var_{Y|X}[Y|X]) = \underbrace{E_X[E_{Y|X}(Y^2|X)}_{=E_Y(Y^2)} - (E_{Y|X}(Y|X))^2]$

$\Rightarrow$ ①+② $= E_Y(Y^2) - [E_Y(Y)]^2 = Var(Y)$

e.g. $Y \sim Geometric(p)$ $\quad P(Y=y) = p^{y-1}(1-p)$

[X is 0 if succeed on the first trial (continue), 1 otherwise (fail)]

$$Var_Y(Y) = Var_X(E_{Y|X}(Y|X)) + E_X(Var_{Y|X}(Y|X))$$
$$= Var_X(X + (1-X)[E_Y(Y)+1]) + E_X(0 \cdot X + (1-X)Var(Y))$$
$$= Var_X(-X E_Y(Y) + 1 + E_Y(Y)) + E_X((1-X)Var_Y(Y))$$
$$= (E_Y(Y))^2 \underbrace{Var_X(X)}_{(1-p)p} + Var_Y(Y)\underbrace{E_X(1-X)}_{p}$$

$P(X=1) = 1-p$
$P(X=0) = p$

$$\Rightarrow (1-p)Var_Y(Y) = (1-p)p(E_Y(Y))^2$$
$$Var_Y(Y) = p(E_Y(Y))^2$$
$$= \frac{p}{(1-p)^2}$$

? $E_Y(Y) = \frac{1}{1-p}$