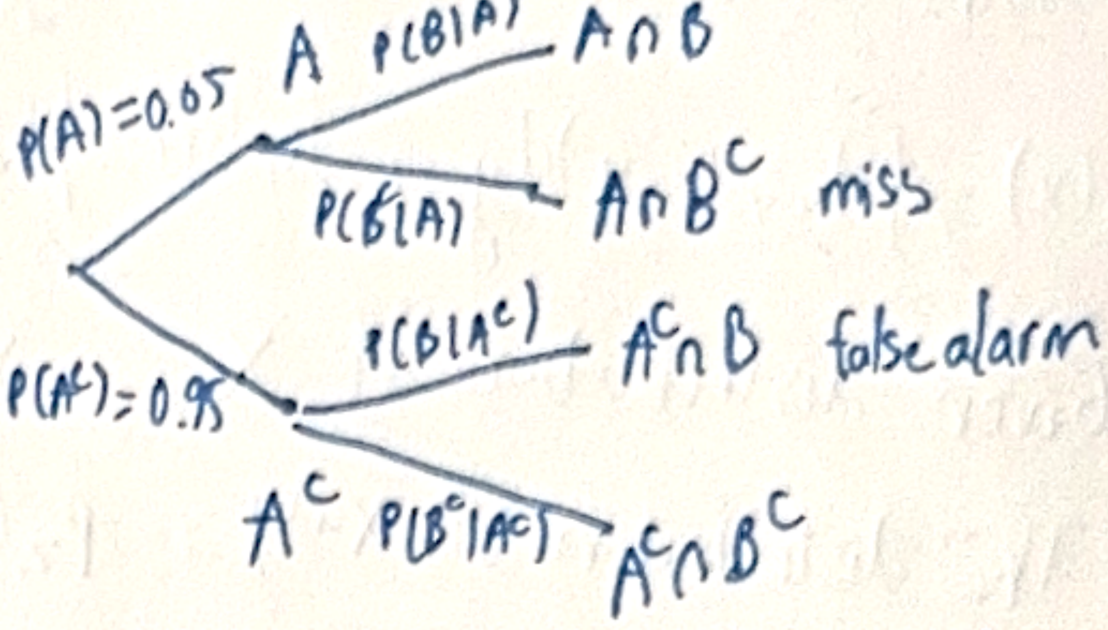
Probability MITx 6431x Đàm Minh Tiến

+ Consequences of the axioms: $P(A\cup B\cup C) = P(A) + P(A^c \cap B) + P(A^c \cap B^c \cap C)$

+ The multiplication rule: $P(A|B) = \dfrac{P(A\cap B)}{P(B)}$; $P(A\cap B) = P(B).P(A|B) = P(A).P(B|A)$



+ Total prob theorem: $P(B) = \sum P(A_i).P(B|A_i)$

+ $P(A_i|B) = \dfrac{P(A_i\cap B)}{P(B)} = \dfrac{P(A_i).P(B|A_i)}{\sum_j P(A_j)P(B|A_j)}$

+ Independence: $P(A\cap B) = P(A).P(B)$ $P(X\le x, Y\le y) = P(X\le x).P(Y\le y)$.

$f_{X,Y}(x,y) = f_X(x).f_Y(y)$ $\Rightarrow E(XY) = E[X].E(Y)$; $V(X+Y) = V(X)+V(Y)$

$P(A|B) = P(A)$ if A and B are independent; $Cov(X,Y)=0 \rightarrow$ uncorrelated, not necessarily dependence

+ Conditional Independence: $P(A\cap B|C) = P(A|C).P(B|C)$ $\rightarrow$ A, B are conditionally indep given C

+ Pairwise Independence does not imply mutual independence.

+ Disjoint events with non-zero prob are not independence ; zero covariance does not imply independence

+ Discrete r.v. / Continuous $E[X] = \sum_x x\,p_X(x)$; $E[g(X)] = \sum g(x).p_X(x)$; $V(X) = E[(X-\mu)^2] = E[X^2] - E[X]^2$

$= \int_{-\infty}^{\infty} x.f_X(x)\,dx$ $V(aX+b) = a^2 V(X)$ $= \int_{-\infty}^{\infty}(x-\mu)^2 f_X(x)\,dx$

+ Moments: n-th moment: $E[X^n] = \begin{cases} \sum_x x^n p_X(x) & \text{(discrete)} \\ \int_{-\infty}^{\infty} x^n.f_X(x)\,dx & \text{(continuous)}\end{cases}$; central moment: $\mu_n' = E[(X-E(X))^n]$

moment generating function: $M_X(t) = E[e^{tX}]$; $E[X^n] = M_X^{(n)}(0)$

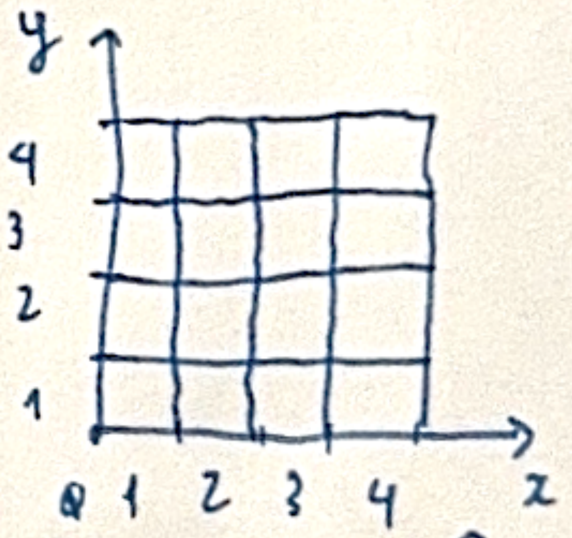
+ Common discrete r.v: $\rightarrow$ binomial distribution $\rightarrow$ approx. Poisson: large n, small p $\rightarrow \lambda = np$

$\hookrightarrow$ approx Normal: large n $\rightarrow \mathcal{N}(np, npq)$

+ Memorylessness property: $P(X>s+t | X>s) = P(X>t)$

+ Joint PMF: $p_{X,Y}(x,y) = P(X=x \text{ and } Y=y)$

Conditional $p_{X|Y}(x|y) = \dfrac{p_{X,Y}(x,y)}{p_Y(y)}$



Independence: X, Y are independence if $p_{X,Y}(x,y) = p_X(x).p_Y(y)$

+ Prob mass function: $p_X(x) = P(X=x)$; $p_X(x) \ge 0$; $\sum_x p_X(x) = 1$ } discrete r.v

PMF $P(a\le X\le b) = \sum_{a\le x\le b} p_X(x)$

+ Prob density function: $P(a\le X\le b) = \int_a^b f_X(x)\,dx$; $f_X(x)\ge 0$; $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$ } continuous r.v

PDF

$P(a\le X\le a+\delta) \approx f_X(a).\delta$; $P(X=a) = 0$

+ Cumulative distribution function: $F_X(x) = P(X\le x) = \int_{-\infty}^{x} f_X(t)\,dt$

CDF

non-decreasing

$\lim_{x\to-\infty} F_X(x) = 0$

$\lim_{x\to+\infty} F_X(x) = 1$

$P(a<X\le b) = F_X(b) - F_X(a)$

+ linear functions of normal r.v. Let $Y = aX+b$, $X\sim\mathcal{N}(\mu,\sigma^2) \rightarrow Y\sim\mathcal{N}(a\mu+b, a^2\sigma^2)$

+ Standardizing a r.v. $X\sim\mathcal{N}(\mu,\sigma^2)$; let $Y = \dfrac{X-\mu}{\sigma}$ $\rightarrow Y\sim\mathcal{N}(0,1)$; $X\sim\mathcal{N}(6,4)$

+ Conditional PDF of X given that $X\in A$: $f_{X|X\in A}(x) = \begin{cases} 0 & \text{, if } x\notin A \\ \dfrac{f_X(x)}{P(A)} & \text{, if } x\in A\end{cases}$

$P(2\le X\le 8) = P\left(\dfrac{2-6}{2}\le\dfrac{X-6}{2}\le\dfrac{8-6}{2}\right)$

$= P(-2\le Y\le 1)$

+ Conditional expectation of X, given that $x \in A$

$$E[X|A] = \sum_x x \, p_{X|A}(x) \qquad E[X|A] = \int x \cdot f_{X|A}(x) \, dx$$

+ Memorylessness of exponential PDF: $P(T > x) = e^{-\lambda x}$ $\qquad P(X > x \mid T > t) = e^{-\lambda x}$

+ Joint continuous r.v and joint PDFs: $f_{X,Y}(x,y) = \dfrac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y) \geq 0$

+ $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy$; $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx$; $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) \, dx \, dy = 1$

+ Given distribution of X; $Y = aX + b \rightarrow p_Y = p_X\left(\frac{y-b}{a}\right)$; $f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$

+ The distribution of $X+Y$: $p_Z(z) = \sum_x p_X(x) p_Y(z-x)$; $f_Z(x) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx$

+ $\text{cov}(X,Y) = E[(X - E[X]).(Y - E[Y])] = E[XY] - E[X]E[Y]$

+ $\text{var}(X_1 + X_2) = \text{var}(X_1) + \text{var}(X_2) + 2\,\text{cov}(X_1, X_2)$

+ correlation coefficient: $\rho(X,Y) = E\left[\frac{(X - E[X])}{\sigma_X} \cdot \frac{(Y - E[Y])}{\sigma_Y}\right] = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y}$

+ law of iterated expectations: $E[E[X|Y]] = E[X]$ $\qquad$ + Markov inequality: $X \geq 0, a \geq 0$; $P(X \geq a) \leq \frac{E[X}{a}$

+ Chebyshev inequality: $P(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2}$

+ CLT: Let $X_1, X_2, \ldots X_n$ be a sequence of i.i.d r.v. with: $E[X_i] = \mu$, $\text{Var}[X_i] = \sigma^2$

$S_n = X_1 + X_2 + .. + X_n$; $\bar{X}_n = \frac{1}{n} . S_n \rightarrow Z_n = \frac{S_n - n\mu}{\sigma \sqrt{n}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$

Then, as $n \rightarrow \infty$, $Z_n \rightarrow \mathcal{N}(0,1)$

+ Confident interval (CI): CI for mean (σ known): $CI = \bar{x} \pm z_{\alpha/2} \left(\frac{\sigma}{\sqrt{n}}\right)$

CI for mean (σ unknown): $CI = \bar{x} \pm t_{\alpha/2, n-1} \left(\frac{s}{\sqrt{n}}\right)$; t-values depend on degrees of freedom

CI for proportion: $CI = \hat{p} \pm z_{\alpha/2} \left(\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$