Lecture 20

Def: If $cov[X,Y] > 0$ we say that the random variables are positively correlated.

If $cov[X,Y] < 0$ then they are negatively correlated.

- Suppose we are interested in a r.v. Y, but we can only have info about a r.v. X. Then a predictor is $\mathbb{E}[(Y-g(X))^2]$

↳ This measures how far the predicted value $g(X)$ is from the actual value Y.

↳ A predictor $g(X)$ is called linear if it has the forum $g(X) = mX + b$, $m,b \in \mathbb{R}$

↳ The best linear predictor of Y based on X is $L_Y(X) = \mu_Y + \frac{\rho \sigma_Y}{\sigma_X}(X - \mu_X)$,

where ρ = corr. coef. XY

$\mu_Y = \mathbb{E}[Y]$

$\mu_X = \mathbb{E}[X]$

↳ Loosely speaking $L_Y(X)$ is the best linear approximation of Y given X

Ex Let $X \sim Bin(n,p)$

$X \sim X_1 + X_2 + \ldots + X_n$

$X_i \sim Ber(p)$ - independent

$\mathbb{E}[X_i] = p$

$Var[X_i] = p(1-p)$

$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = np$

$\overset{\text{indep.}}{\Longrightarrow} Var[X] = \sum_{i=1}^{n} Var[X_i] = np(1-p)$

Multi-Dimensional discrete random vectors

Def: Suppose that $X_1, \dots, X_n$ are discrete r.v. with ranges $\mathcal{X}_1, \dots, \mathcal{X}_n$ and pmf's $p_1, \dots, p_n$. The joint pmf of the n-dim random vector $(X_1, \dots, X_n)$ is the function $p: \mathcal{X}_1 \cdot \dots \cdot \mathcal{X}_n \to [0,1]$ defined by:

$$p(x_1, \dots, x_n) = \mathbb{P}(X_1 = x_1, \dots, X_n = x_n)$$

Cor: Suppose that random variables $X_1, \dots, X_n$ are <u>independent</u>. Then the following holds:

i) $\forall$ $1 \leq k \leq n-1$ and any functions $f(x_1, \dots, x_n)$ and $g(x_1, \dots, x_n)$ the random variables $f(X_1, \dots, X_k) \perp\!\!\!\perp g(x_1, \dots, X_n)$

ii) For any functions $f(x_1), \dots, f(x_n)$ the random variable $f_1(x_1), \dots, f_n(X_n)$ are independent.

Prop: Law of Subconscious Statistician

↳ Suppose that $X_1, \dots, X_n$ are discrete r.v. with ranges $\mathcal{X}_1, \dots, \mathcal{X}_n$ and joint pmf $p(x_1, \dots, x_n)$

i) If $\mathcal{X}_1 \times \dots \times \mathcal{X}_n \to \mathbb{R}$ is a function, then

$$\mathbb{E}[f(X_1 \dots X_n)] = \sum_{(x_1, \dots x_n) \in \mathcal{X}_1 \times \dots \times \mathcal{X}_n} f(x_1, \dots, x_n) \, p(x_1, \dots, x_n)$$

ii) If the random variables $X_1, \dots, X_n$ are independent, then for any functions $f_k : X_k \to \mathbb{R}$, $k = 1, \dots, n$ The rand variables $f_1(X_1), \dots, f_n(X_n)$ are indep.

and we have:

$$\mathbb{E}\left[\prod_{k=1}^{n} f_k(X_k)\right] = \prod_{k=1}^{n} \mathbb{E}[f_k(X_k)]$$

X - Y - random variables

$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ TRUE

$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ FALSE, need uncorrelation (which implies independence)

$\mathrm{var}[X+Y] = \mathrm{var}[X] + \mathrm{var}[Y] \Rightarrow \mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ TRUE

$\rho(X,Y) = 0 \implies p(x,y) = \underbrace{p_x(x)\, p_y(y)}_{\text{marginals}}$ FALSE, if you flip arrow then true