more joint distributions, conditional distributions, tail bounds

LECTURE 17

JOINT PROBABILITIES

If X and Y are discrete....

$\mathbb{P}(a \le X \le b \cap c \le Y \le d) = \sum_{x \in a \le X \le b} \sum_{y \in c \le Y \le d} p_{X,Y}(x,y)$

sum over the joint PMF for all pairs of x and y that fall in this range

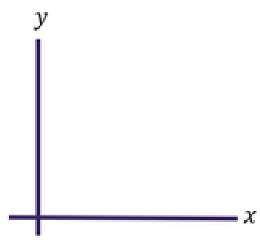
If X and Y are continuous...

$\mathbb{P}(a \le X \le b \cap c \le Y \le d) = \int_a^b \int_c^d f_{X,Y}(x,y)\,dy\,dx$

integrate over the joint PDF for all pairs of x and y that fall in this range

The time until server 1 crashes is X~Exp(u), and the time until server 2 crashes is Y~Exp(v). Both servers are independent of each other. What is the probability server 1 crashes before server 2?

P(X<Y) =



sketch the region to integrate over

COVARIANCE

$\text{Cov}(X,Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$

Measures dependence between X and Y:

> Covariance is **positive** -> they are *positively correlated* (If X increases, Y tends to also increase

> Covariance is **negative** -> they are *negatively correlated* (If X increases, Y tends to decrease)

> $\text{Cov}(X,Y) = \text{Cov}(Y,X)$

> $\text{Cov}(X,X) = \text{Var}(X)$
when you plug in X, X, we get $\mathbb{E}[(X - \mathbb{E}[X])^2]$, the variance

> $\text{Cov}(aX + b, Y) = a \cdot \text{Cov}(X,Y)$
linearity of expectation

> $\text{Cov}(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j) = \sum_{i=1}^{n}\sum_{j=1}^{m} \text{Cov}(X_i, Y_j)$

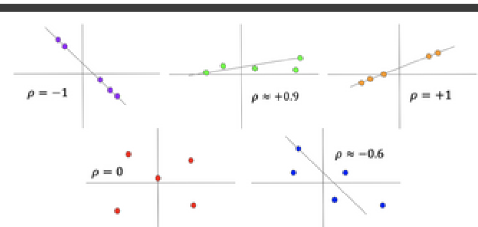
VARIANCE OF SUM OF RANDOM VARIABLES WHEN THEY'RE NOT INDEPENDENT

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

You and your friend are playing a game, you flip a coin: if heads you pay your friend a dollar, if tails they pay you a dollar. Let X be your profit and Y be your friend's profit. What is Var(X+Y)?

PEARSON CORRELATION (NORMALIZED COVARIANCE)

Pearson correlation: $\rho(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma_X \cdot \sigma_Y} = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}}$



CONDITIONAL DISTRIBUTIONS

CONDITIONAL PMF / PDF

Conditional PMF: $p_{X|Y}(a|b) = \mathbb{P}(X = a|Y = b) = \frac{p_{X,Y}(a,b)}{p_Y(b)} = \frac{p_{Y|X}(b,a)\, p_X(a)}{p_Y(b)}$

Conditional PDF: $f_{X|Y}(a|b) = \frac{f_{X,Y}(a,b)}{f_Y(b)} = \frac{f_{Y|X}(b,a)\, f_X(a)}{f_Y(b)}$

CONDITIONAL EXPECTATION

we can define things like "**conditional expectations**" which is the expectation of a random variable in the new probability space. *Note how the continous case is the same but sum->integral, PMF->density*

$$\mathbb{E}[X|A] = \sum_{k \in \Omega} k \cdot \mathbb{P}(X = k|A)$$

$\mathbb{E}[X|Y = y] = \sum_{k \in \Omega_X} k \cdot \mathbb{P}(X = k|Y = y)$

or $\mathbb{E}[X|Y = y] = \int_{-\infty}^{\infty} k \cdot f_{X|Y}(k, y)\, dk$ if continuous

LAW OF TOTAL EXPECTATION

we can define things like "**conditional expectations**" which is the expectation of a random variable in the new probability space. *Note how the continous case is the same but sum->integral, PMF->density*

$$\mathbb{E}[X|A] = \sum_{k \in \Omega} k \cdot \mathbb{P}(X = k|A)$$

$\mathbb{E}[X|Y = y] = \sum_{k \in \Omega_X} k \cdot \mathbb{P}(X = k|Y = y)$

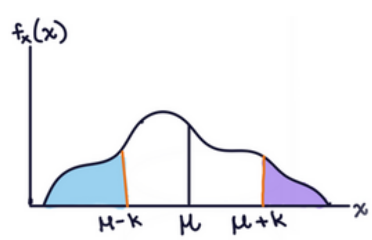
or $\mathbb{E}[X|Y = y] = \int_{-\infty}^{\infty} k \cdot f_{X|Y}(k, y)\, dk$ if continuous

we can define things like "**conditional expectations**" which is the expectation of a random variable in the new probability space. *Note how the continous case is the same but sum->integral, PMF->density*

TAIL BOUNDS

WHAT ARE TAIL BOUNDS?

A tail bound (or concentration inequality) is a statement bounds the probability in the "tails" of the distribution (e.g., there's little probability far from the center) or (equivalently) the probability is concentrated near the expectation.



MARKOV'S INEQUALITY

Let X be a random variable supported (only) on non-negative numbers. For any $t > 0$

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t}$$

Requirements:

1. X must be non-negative
2. We know the expectation of X

You roll a fair die until you see a 6. Let X be the number of rolls. Bound P(X≥12.)

The average number of ads on a website is 25. Upper bound the prob. of a website with >= 75 ads.