**The Austalian National University**
**School of Computing**
Jo Ciucă & Thang Bui

Semester 2, 2023
Tutorial 7

# COMP3670/6670: Introduction to Machine Learning

## Question 1 — Marginals and Conditionals

Consider two discrete random variables $X$ and $Y$ , with sample spaces $\{1, 2, 3\}$ and $\{1, 2\}$ respectively, and the following joint distribution $p(x, y)$.

| $p(x,y)$ | $x = 1$ | $x = 2$ | $x = 3$ |
|---|---|---|---|
| $y = 1$ | 1/16 | 4/16 | 1/16 |
| $y = 2$ | 2/16 | 3/16 | 5/16 |

1. Compute the marginal distributions $p(x)$ and $p(y)$
2. Compute the conditional distributions $p(x \mid Y = 1)$ and $p(y \mid X = 2)$.
3. Are $X$ and $Y$ statistically independent?
4. Compute the expectation values $\mathbb{E}[X]$ and $\mathbb{E}[Y]$. [1]

## Question 2 — Bayes' Rule

Here is a bag containing three coins: A fair coin (equally likely to land on heads or tails), a two headed coin (always lands on heads) and a two tailed coin (always lands on tails).

1. You select one of the coins uniformly at random, and flip it. The result is heads. What is the probability of the other side being tails?
2. You select one of the coins uniformly at random, and flip the coin $N$ times. The outcome of every trial is that the coin lands on heads. What is the probability of the other side being tails? What does this result tend to as $N \to \infty$? Can you explain the result?

## Question 3 — Expected Value

A crooked gambler approaches you with the opportunity to play a game.
"I've got a perfectly ordinary deck of 52 cards. It costs \$1 to play! I draw three cards from the deck. If two of them are red, you win \$1. If all three are red, you win five dollars! What do you say?"

1. Should you play this game or not?
2. You reply to the gambler "I'll play if we change the rules, and each time you draw a card, we write the result down, and then reshuffle the card back into the deck."
   Should you play this game now?

## Question 4 — Properties of Conditional Distributions

Let $X, Y$ be random variables, with corresponding probability distribution functions $p(x) : \mathbb{R} \to \mathbb{R}$ and $p(y) : \mathbb{R} \to \mathbb{R}$ respectively. Let $p(x, y) : \mathbb{R}^2 \to \mathbb{R}$ be the joint probability distribution function. We define the expectation value of a function $f : \mathbb{R} \to \mathbb{R}$ of $X$ to be

$$\mathbb{E}_X[f(x)] = \int_{-\infty}^{\infty} f(x)p(x)dx$$

[1] Note that $\mathbb{E}[X]$ is shorthand for $\mathbb{E}_X[x]$.

and the expectation value of a function $g : \mathbb{R}^2 \to \mathbb{R}$ with respect to $X$ and $Y$ to be

$$\mathbb{E}_{X,Y}[g(x,y)] := \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)p(x,y)dxdy$$

1. Prove that if a binary function $g(x,y) : \mathbb{R}^2 \to \mathbb{R}$ has no dependence on the second argument $y$ (that is, if $g(x,y) = h(x)$ for some function $h : \mathbb{R} \to \mathbb{R}$), then

$$\mathbb{E}_{X,Y}[g(x,y)] = \mathbb{E}_X[h(x)]$$

2. It was given in lectures that the covariance between two random variables can be expressed in one of two ways:

$$\mathrm{Cov}_{X,Y}[x,y] := \mathbb{E}_{X,Y}\left[(x - \mathbb{E}_X[x])(y - \mathbb{E}_Y[y])\right]$$

   or as the alternate form

$$\mathrm{Cov}_{X,Y}[x,y] = \mathbb{E}_{X,Y}[xy] - \mathbb{E}_X[x]\mathbb{E}_Y[y]$$

   Prove these are equivalent.

3. It was given in lectures that if $X$ and $Y$ are statistically independent, then $\mathrm{Cov}_{X,Y}[x,y] = 0$ Prove this.