Lecture 8, Jan 25, 2023

Joint Probability Distributions (Discrete)

- Often events are correlated, and we want to look at the probability of two events together

Definition

$f(x, y)$ is the *joint probability mass function* of the discrete varaibles X and Y if:
- $f(x, y) \geq 0, \forall (x, y) \in S$
- $\displaystyle\sum_{(x,y)\in S} f(x, y) = 1$
- $f(x, y) = P(X = x, Y = y)$

- Example: Drawing 5 cards from a deck of 52, X is the number of queens and Y is the number of kings, what is $f(x, y)$?
  - Total number of hands is $\binom{52}{5}$
  - Number of hands with x queens and y kings is $\binom{4}{x}\binom{4}{y}\binom{52-8}{5-x-y}$ for $x, y \leq 4, x + y \leq 5$
  - Therefore $f(x, y) = \begin{cases} 0 & x + y > 5, x > 4, y > 4 \\ \dfrac{\binom{4}{x}\binom{4}{y}\binom{44}{5-x-y}}{\binom{52}{5}} & x + y \leq 5, x \leq 4, y \leq 4 \end{cases}$
- Continued example: Consider $A = \{\, (x + y) \mid x + y = 2 \,\}$; then $P((x, y) \in A) = \displaystyle\sum_{(x,y)\in A} f(x, y) = f(0, 2) + f(1, 1) + f(2, 0)$

Joint Probability Distributions (Continuous)

Definition

$f(x, y)$ is the *joint probability density function* of the continuous variables X and Y if
- $f(x, y) \geq 0, \forall (x, y) \in S$
- $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = 1$
- If $A \in S$, $P((x, y) \in A) = \displaystyle\int_{(x,y)\in A} f(x, y)\,\mathrm{d}x\,\mathrm{d}y$

- Example: Uniform distribution, $S = \{\, (x, y) \mid 1 \leq x \leq 1, -1 \leq y \leq 1 \,\}$
  - A uniform distribution means the probability of any outcome is the same
  - $f(x, y) = \dfrac{1}{4}$ for $(x, y) \in S$, since the "area" is 4
  - Consider an event $A = (x, y)|0 \leq x \leq 1, 0 \leq y \leq 1$, now $P((x, y) \in A) = \dfrac{1}{4}$ since A takes up a quarter of S
    - We can also use $\displaystyle\int_0^1\int_0^1 \frac{1}{4}\,\mathrm{d}x\,\mathrm{d}y = \frac{1}{4}$

Marginal Distributions

- Suppose we know $f(x, y)$ for X, Y; we can find the distribution for X as $g(x) = \displaystyle\sum_y f(x, y)$ in the discrete case, and $g(x) = \displaystyle\int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}y$ in the continuous case

  - This is known as a *marginal distribution*
- Example: Given a PDF $f(x, y) = \begin{cases} 1 \\ |x| + |y| \leq 1, y \geq 0 \end{cases}$ 0otherwise, find $g(x)$
  - This describes a triangle above the x axis
  - $g(x) = \int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}y \int_{0}^{1-|x|} 1\,\mathrm{d}y = 1 - |x|$