Math 525

Yiwei Fu

Spet 20

1 Random Walk

Start with Problem 3.11.3

$$\mathbb{E}(f(X)) = \sum_x f(x) p_X(x).$$

It is also true for jointly discrete RV's.

Can use this for our argument about "probabilistic method" for counting

Theorem 1.1. *Suppose $X_1, \dots, X_n$ are jointly distributed RV's (discrete), then*

$$\mathbb{E}(X_1 + \ldots + X_n) = \mathbb{E}(X_1) + \ldots + \mathbb{E}(X_n).$$

Proof.

$$\begin{aligned}
\mathbb{E}(f) &= \mathbb{E}(X_1 + \ldots X_n) \\
&= \sum_x (x_1 + x_2 + \ldots + x_n) p_{x_1 \ldots x_n}(x_1, \dots, x_n) \\
&= \sum_i \sum_x x_i p_{x_1 \ldots x_n}(x_1, \dots, x_n).
\end{aligned}$$

e.g. $i = 1 : \sum_x x_1 p_{x_1 \ldots x_n}(x_1, \dots, x_n)$

$$x = (x_1, x_2, \dots, x_n) = (x_1, x')$$

$$\begin{aligned}
\sum_x x_1 p_{x_1 \ldots x_n}(x_1, \dots, x_n) &= \sum_{x_1} x_1 \sum_{x'} p_{x_1 \ldots x_n}(x_1, x') \\
&= \sum_{x_1} x_1 P_{X_1}(x_1) = \mathbb{E}(X_1).
\end{aligned}$$

■

Notice: X_i need not to be independent.

For Canvas, in "File", there are handbooks of distributions.

Conditional Expectation:

Suppose X, Y are jointly distributed RV's. Then conditional expectation:

$$\mathbb{E}(Y \mid X = x) = \sum_{y_i} y_i \mathbb{P}(Y = y_i \mid X = x) = \sum_{y_i} \frac{p_{X,Y}(y_i, x)}{p_X(x)}.$$

Notice:

1. If X, Y are independent then
$$\mathbb{E}(Y \mid X = x) = \mathbb{E}(Y).$$

2. $\mathbb{E}(Y \mid X = x)$ in general an RV $= f(X) = \mathbb{E}(Y \mid X)$. We have
$$\mathbb{E}(\mathbb{E}(Y \mid X)) = \mathbb{E}(Y).$$

$$\begin{aligned}
\mathbb{E}(\mathbb{E}(Y \mid X)) &= \sum_x \mathbb{E}(Y \mid X = x) p_X(x) \\
&= \sum_x \sum_y \mathbb{P}(Y = y \mid X = x) p_X(x) \\
&= \sum_{x,y} y \frac{p_{X,Y}(y, x)}{p_X(x)} p_X(x) \\
&= \sum_y y \sum_x p_{X,Y}(y, x) \\
&= \sum_y y p_Y(y) \\
&= \mathbb{E}(Y).
\end{aligned}$$

"Moments" $\quad \mathbb{E}(X^n)$.

"1st moment"$= \mathbb{E}(X^1) = \mathbb{E}(X)$.

"2nd moment"$= \mathbb{E}(X^2) \geq 0$.

Why study moments?

Consider continuous distribution on $[0, 1]$. We have a p.d.f $f_X(x)$.

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x) dx = \int_0^1 \chi_{[a,b]}(x) f_X(x) dx$$

where $\chi_{[a,b]}(x)$ is the characteristic function

$$\chi_{[a,b]}(x) = \begin{cases} 1, & x \in [a, b] \\ 0, & \text{otherwise.} \end{cases}$$

Moments:

$$\int_0^1 x^n f_X(x) dx \Rightarrow \int_0^1 P(X) f(x) dx$$

Weierstrass approximation theorem: $g(x)$ is continuous on $[0,1]$ there there exists a polynomial $P_i(x)$ $s.t.$ $|g(x) - P_i(x)| < \frac{1}{i}$ for $i >> 0$.

Invariant of a distribution:

What is the difference between

$$X = \begin{cases} 1, & \frac{1}{2} \\ -1, & \frac{1}{2} \end{cases}, \; Y = \begin{cases} 100, & \frac{1}{2} \\ -100, & \frac{1}{2} \end{cases}.$$

We have

$$\mathbb{E}(X) = \mathbb{E}(Y) = 0.$$

Spread:

$$\mathbb{E}(X^2) = 1, \mathbb{E}(Y^2) = 10000.$$

In general, $\mathrm{Var}(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$.

Notice:

$$\mathrm{Var}(X) = 0 \iff \mathrm{Prob}(X = \mathbb{E}(X)) = 1$$

Proof. Denote $\mathbb{E}(X) = \mu$.

$$\begin{aligned} \mathrm{Var}(X) &= \mathbb{E}((X - \mathbb{E}(X))^2) \\ &= \sum_x (x - \mu)^2 p_X(x) \\ &= 0 \implies p_X(x) = 0 \text{ if } x - \mu \neq 0. \end{aligned}$$

■