Discrete Random Variables

Chapter 3: Discrete Random Variables

Anduo Wang
Temple University
Email: anduo.wang@gmail.com
https://cis-linux1.temple.edu/∼tug29203/25fall-2033/index.html

Independent Random Variables

- The concept of independent random variables is very similar to independent events.
- A and B are independent if $P(A, B) = P(A)P(B)$.
- We have the definition of independent random variables

> Consider two discrete random variables X and Y. We say that X and Y are independent if
>
> $$P\Big(X = x, Y = y\Big) = P(X = x)P(Y = y), \forall \text{ x,y.}$$
>
> In general, if two random variables are independent, then you can write
>
> $$P\Big(X \in A, Y \in B\Big) = P(X \in A)P(Y \in B).$$

Independent Random Variables

- Intuitively, two random variables X and Y are independent if knowing the value of one does not change the probabilities for the other one, i.e.

$$P(Y = y|X = x) = P(Y = y), \forall \text{ x,y}$$

- In general

Consider n discrete random variables $X_1, X_2, X_3, ..., X_n$. We say that $X_1, X_2, X_3, ..., X_n$ are independent if

$$P\Big(X_1 = x_1, X_2 = x_2, ..., X_n = x_n\Big)$$
$$= P(X_1 = x_1)P(X_2 = x_2)...P(X_n = x_n)$$

for all $x_1, x_2, ..., x_n$

Independent Random Variables

- It is possible to argue that two random variables are independent simply because they do not have any physical interactions with each other.
- Example: I toss a coin and define X to be the number of heads I observe. Then I toss the coin two more times and define Y to be the number of heads that I observe this time.
- What is probability $P((X < 2) \cap (Y > 1))$?

Expectation and Variance

- If we want to summarize a random variable by a single number, then this number should undoubtedly be its *expected value*.
  - The *expectation* or *mean*, gives the *center* — in the sense of average value — of the distribution of the random variable.
- A second number to describe the random variable, then we look at its *variance*
  - A measure of spread of the distribution of the random variable.

Motivating example: expected values

An oil company needs drill bits in an exploration project. Suppose that it is known that (after rounding to the nearest hour) drill bits of the type used in this particular project will last 2, 3, or 4 hours with probabilities 0.1, 0.7, and 0.2. If a drill bit is replaced by one of the same type each time it has worn out, how long could exploration be continued if in total the company would reserve 10 drill bits for the exploration job?

- Answer: the weighted average: $0.1 \cdot 2 + 0.7 \cdot 3 + 0.2 \cdot 4 = 3.1$, so the exploration could continue for 10×3.1 hours.

Motivating example: expected values

An oil company needs drill bits in an exploration project. Suppose that it is known that (after rounding to the nearest hour) drill bits of the type used in this particular project will last 2, 3, or 4 hours with probabilities 0.1, 0.7, and 0.2. If a drill bit is replaced by one of the same type each time it has worn out, how long could exploration be continued if in total the company would reserve 10 drill bits for the exploration job?

- Answer: the weighted average: $0.1 \cdot 2 + 0.7 \cdot 3 + 0.2 \cdot 4 = 3.1$, so the exploration could continue for 10×3.1 hours.
- The weighted average is the **expected value or expectation** of the random variable X whose distribution is given by

$$P(X = 2) = 0.1, P(X = 3) = 0.7, P(X = 4) = 0.2.$$

Motivating example: expected values

The conclusion about a 31-hour total drilling time is correct in the following sense: for a large number n of drill bits the total running time will be around n times 3.1 hours with high probability. In the example, where $n = 10$, we have, for instance, that drilling will continue for 29, 30, 31, 32, or 33 hours with probability more than 0.86, while the probability that it will last only for 20, 21, 22, 23, or 24 hours is less than 0.00006.

- Later: the law of large numbers ...

Expectation

Definition

The expectation of a discrete random variable X taking the values $a_1, a_2, \cdots$ and with probability mass function p is the number

$$E[X] = \Sigma_i a_i P(X = a_i) = \Sigma_i a_i p(a_i).$$

Expectation

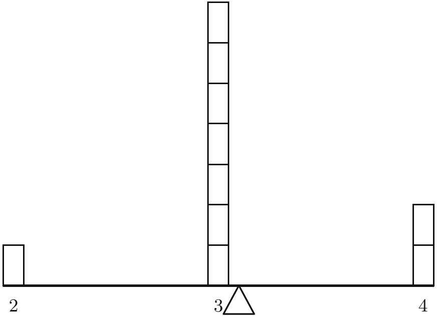
- The average of a random variable X is otherwise called its **expected value** or **mean**.
- The expected value is defined as the weighted average of the values in the range.

Expected value of X: $EX = E[X] = E(X) = \mu_X$ **Definition:** Let X be a discrete random variable with range $R_X = \{x_1, x_2, x_3, \ldots\}$ (finite or countably infinite). The *expected* value of X, is defined as

$$EX = \sum_{x_k \in R_X} x_k \mathrm{P}(X = x_k) = \sum_{x_k \in R_X} x_k P_X(x_k).$$

Expectation as a weighted average: illustration

a more physical interpretation of this notion, namely as the center of gravity of weights $p(a_i)$ placed at the points a_i, for the random variable associated with the drill bit



The geometric distribution

Examples

If you buy a lottery ticket every week and you have a chance of 1 in 10 000 of winning the jackpot, what is the expected number of weeks you have to buy tickets before you get the jackpot?

The geometric distribution

Examples

If you buy a lottery ticket every week and you have a chance of 1 in 10 000 of winning the jackpot, what is the expected number of weeks you have to buy tickets before you get the jackpot?

- Ans: 10000 weeks (almost two centuries!)
- The number of weeks is modeled by a random variable with a geometric distribution with parameter $p = 10^{-4}$

The geometric distribution

Examples

If you buy a lottery ticket every week and you have a chance of 1 in 10 000 of winning the jackpot, what is the expected number of weeks you have to buy tickets before you get the jackpot?

- Ans: 10000 weeks (almost two centuries!)
- The number of weeks is modeled by a random variable with a geometric distribution with parameter $p = 10^{-4}$

The expectation of a geometric distribution.

Let X have a geometric distribution with parameter p; then

$$E[X] = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = \frac{1}{p}.$$

Expectation

- Expectation is linear:

Definition: We have

- $E[aX + b] = aEX + b$, for all $a, b \in \mathbb{R}$;
- $E[X_1 + X_2 + \cdots + X_n] = EX_1 + EX_2 + \cdots + EX_n$, for any set of random variables $X_1, X_2, \cdots, X_n$.

- We can use the linearity of expectation to easily calculate the expected value of binomial and Pascal distributions.

Expectation

- The expectation for different special distributions
  - $X \sim \text{Bernoulli}(p)$: $EX = p$.
  - $X \sim \text{Geometric}(p)$: $EX = \frac{1}{p}$.
  - $X \sim \text{Poisson}(\lambda)$: $EX = \lambda$.
  - $X \sim \text{Binomial}(n, p)$: $EX = np$.
  - $X \sim \text{Pascal}(m, p)$: $EX = \frac{m}{p}$.
  - $X \sim \text{Hypergeometric}(b, r, k)$: $EX = \frac{kb}{b+r}$.

Functions of random variables

- Often one does not want to compute the expected value of a random variable X but rather of a function of X

Examples

Suppose an architect wants maximal variety in the sizes of buildings: these should be of the same width and depth X. Let X be a discrete random variable with $P_X(k) = \frac{1}{5}$ for $k = 1, 3, 5, 9, 10$ meters. What is the distribution of the area $Y = X^2$ of a building?

- Let us compute F_Y, for $a = 1, 9, 25, 81, 100$:
  $F_Y(a) = P(X^2 \leq a) = P(X = \sqrt{a}) = 1/5$

Functions of Random Variables

- If X is a random variable and $Y = g(X)$, then Y itself is a random variable.
- We can thus define its PMF,CDF and expected value.
- The range of Y, R_Y is given as

$$R_Y = \{g(x) | x \in R_X\}$$

- If we knew the PMF of X, we can obtain the PMF of Y as

$$\begin{aligned} P_Y(y) &= P(Y = y) \\ &= P(g(X) = y) \\ &= \sum_{x:g(x)=y} P_X(x). \end{aligned}$$

Functions of Random Variables

- Example: Let X be a discrete random variable with $P_X(k) = \frac{1}{5}$ for $k = -1, 0, 1, 2, 3$. Let $Y = 2|X|$.
- Find the range and PMF of Y.

Expected Value of a Function of a Random Variable - LOTUS

- Let X be a discrete random variable and $Y = g(X)$.
- To calculate the expected value of Y, we can use LOTUS.

Law of the unconscious statistician (LOTUS) for discrete random variables

If X is discrete, taking the values $a_1, a_2, \cdots$, then

$$E(g(X)) = \sum_i g(a_i)P(X = a_i).$$

Expected Value of a Function of a Random Variable - LOTUS

- Let X be a discrete random variable and $Y = g(X)$.
- To calculate the expected value of Y, we can use LOTUS.

Law of the unconscious statistician (LOTUS) for discrete random variables

If X is discrete, taking the values $a_1, a_2, \cdots$, then

$$E(g(X)) = \sum_i g(a_i)P(X = a_i).$$

- Prove: expectation is linear

Expected Value of a Function of a Random Variable - LOTUS

- Let X be a discrete random variable with range $R_X = \{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}, \pi\}$, such that $P_X(0) = P_X(\frac{\pi}{4}) = P_X(\frac{\pi}{2}) = P_X(\frac{3\pi}{4}) = P_X(\pi) = \frac{1}{5}$.
- Find $E[sin(X)]$.

Variance

Suppose you are offered an opportunity for an investment whose expected return is 500. If you are given the extra information that this expected value is the result of a 50% chance of a $450 return and a 50% chance of a 550 return, then you would not hesitate to spend 450 on this investment. However, if the expected return were the result of a 50% chance of a 0 return and a 50% chance of a 1000 return, then most people would be reluctant to spend such an amount.

- The spread (around the mean) of a random variable is of great importance.

Definition

The variance $Var(X)$ of a discrete random variable X is the number.

$$Var(X) = E(X - E[X])^2.$$

Variance

- Consider random variables X and Y with PMFs

$$P_X(x) = \begin{cases} 0.5 & \text{for } x = -100 \\ 0.5 & \text{for } x = 100 \\ 0 & \text{otherwise} \end{cases}$$

$$P_Y(y) = \begin{cases} 1 & \text{for } y = 0 \\ 0 & \text{otherwise} \end{cases}$$

- $EX = EY = 0$ but the two distributions are very different.
- **Variance** is a measure of how spread out the distribution of a random variable is.
- Variance of Y is quite small as the distribution is concentrated, while the variance of X is larger.

Variance

The **variance** of a random variable X, with mean $EX = \mu_X$, is defined as

$$\mathrm{Var}(X) = E\big[(X - \mu_X)^2\big].$$

Computational formula for the variance:

$$\mathrm{Var}(X) = E\left[X^2\right] - \left[EX\right]^2 \tag{1}$$

where $E[X^2] = EX^2 = \sum_{x_k \in R_X} x_k^2 P_X(x_k)$.

The drill bit example revisited

- Recall: X takes the values 2, 3, and 4 with probabilities 0.1,0.7, and 0.2.

$$
\begin{aligned}
Var(X) &= E[(X-3.1)^2] \\
&= 0.1\cdot(2-3.1)^2 + 0.7\cdot(3-3.1)^2 + 0.2\cdot(4-3.1)^2 \\
&= 0.121 + 0.007 + 0.162 \\
&= 0.29 \\
Var(X) &= E[X^2] - (3.1)^2 \\
&= 0.1\cdot 22 + 0.7\cdot 32 + 0.2\cdot 42 - 9.61 \\
&= \cdots
\end{aligned}
$$

Standard Deviation and Useful Results

The **standard deviation** of a random variable X is defined as

$$\mathrm{SD}(X) = \sigma_X = \sqrt{\mathrm{Var}(X)}.$$

Theorem: Variance is not a linear operator. For a random variable X and real numbers a and b,

$$\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$$

Standard Deviation and Useful Results

Theorem: If $X_1, X_2, \cdots, X_n$ are independent random variables and $X = X_1 + X_2 + \cdots + X_n$, then

$$\mathrm{Var}(X) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_n)$$

- Example: If $X \sim Binomial(n, p)$, find $Var(X)$.