NMAI059 Probability and statistics 1
Class 3

Robert Šámal

Overview

Discrete random variables

Examples of discrete r.v.'s

Expectation

Random variable

Often we are interested in a number given as a result of a random experiment.

- We throw a dart and measure the distance from the center of the dartboard.
- We roll a die until we get a six, then count how many rolls it took.
- In a quicksort algorithm (with a random choice of pivot) we measure the number of operations.

Definition

Given a probability space $(\Omega, \mathcal{F}, P)$*. We call a function* $X : \Omega \to \mathbb{R}$ *a discrete random variable, if* $Im(X)$ *(range of* X*) is a countable set and if for every real* x *we have*

$$\{\omega \in \Omega : X(\omega) = x\} \in \mathcal{F}.$$

PMF

Definition

Probability mass function, PMF of a discrete random variable X *is a function* $p_X : \mathbb{R} \to [0,1]$ *such that*

$$p_X(x) = P(X = x) = P(\{\omega \in \Omega : X(\omega) = x\})$$

- $\sum_{x \in Im(X)} p_X(x) =$?
- $S := Im(X) \qquad Q(A) := \sum_{x \in A} p_X(x)$
  $(S, \mathcal{P}(S), Q)$ is a discrete probability space.
- For $S = \{s_i : i \in I\}$ countable set of reals and $c_i \in [0,1]$ satisfying $\sum_{i \in I} c_i = 1$ there is a probability space and a discrete r.v. X on it such that $p_X(s_i) = c_i$ for $i \in I$.

Another description – CDF

Definition

Cumulative distribution function, CDF of a r.v. X is a function

$$F_X(x) := P(X \leq x) = P(\{\omega \in \Omega : X(\omega) \leq x).$$

- F_X is a nondecreasing function
- $\lim_{x\to-\infty} F_X(x) = 0$
- $\lim_{x\to+\infty} F_X(x) = 1$
- F_X is right-continuous

Overview

Discrete random variables

Examples of discrete r.v.'s

Expectation

Bernoulli/alternate distribution

- $X =$ number of tails in one toss of a coin (not necessary a fair one)
- We write $X \sim Bern(p)$. (Sometimes $Alt(p)$.)

- Given $p \in [0, 1]$.
- $p_X(1) = p$
- $p_X(0) = 1 - p$
- $p_X(k) = 0$ for $k \neq 0, 1$

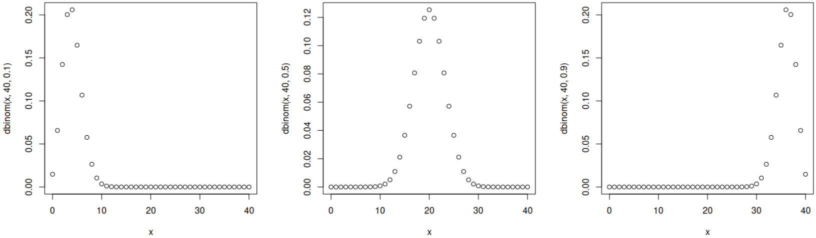
- For an event $A \in \mathcal{F}$ we define *indicator random variable* I_A:
- $I_A(\omega) = 1$ if $\omega \in A$, $I_A(\omega) = 0$ otherwise.
- $I_A \sim Bern(P(A))$

Binomial distribution

- $X =$ number of tails in n independent tosses of a loaded coin.
- Given $p \in [0, 1]$.
- We write $X \sim Bin(n, p)$.

- $X = \sum_{i=1}^{n} X_i$ for independent r.v.'s $X_1, \ldots, X_n \sim Bern(p)$.
- $p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$ for $k \in \{0, 1, \ldots, n\}$

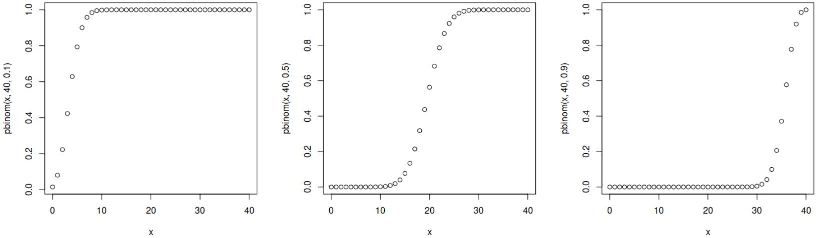
Binomial distribution: PMF



Generated by the following code in R

```
x <- 0:40
plot(x,dbinom(x,40,0.1))
plot(x,dbinom(x,40,0.5))
plot(x,dbinom(x,40,0.9))
```

Binomial distribution: CDF



Generated by the following code in R

```
x <- 0:40
plot(x,pbinom(x,40,0.1))
plot(x,pbinom(x,40,0.5))
plot(x,pbinom(x,40,0.9))
```

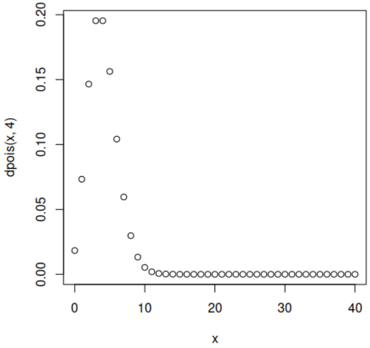
Hypergeometric distribution

- $X =$ the number of red balls we get out of n, when the urn contains K red out of N balls
- Given n, N, K.
- We write $X \sim Hyper(N, K, n)$.
- $p_X(k) = P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$

Poisson distribution

- We write $X \sim Pois(\lambda)$.
- Given real $\lambda > 0$.
- $p_X(k) = \frac{\lambda^k}{k!} e^{-\lambda}$
- $Pois(\lambda)$ is a limit of $Bin(n, \lambda/n)$
- X describes, e.g., the number of emails we get in a day.

Poisson distribution: PMF



Generated by the following code in R

```
x <- seq(0,40,by=1)
plot(x,dpois(x,4))
```

Poisson paradigm

- $A_1, \ldots, A_n$ are (almost-)independent events with $P(A_i) = p_i$, $\lambda = \sum_i p_i$. Suppose n is large, each of p_i small. Then it is approximately true that

$$\sum_{i=1}^{n} I_{A_i} \sim Pois(\lambda).$$

Geometric distribution

- $X =$ number of coin tosses till we get a tail
- We write $X \sim Geom(p)$.

- Given $p \in [0, 1]$.
- $p_X(k) = (1-p)^{k-1}p$, for $k = 1, 2, \ldots$

- Some people call this distribution shifted geometric, the normal geometric would then be distribution of $X - 1$, that is the number of unsuccessful tosses.

Overview

Discrete random variables

Examples of discrete r.v.'s

Expectation

Expectation

Definition

Given a discrete r.v. X, its expectation is denoted by $\mathbb{E}(X)$ and defined by

$$\mathbb{E}(X) = \sum_{x \in Im(X)} x \cdot P(X = x),$$

whenever the sum is defined.

- Suppose X is defined on a discrete space $(\Omega, \mathcal{F}, P)$. Then we can also define the expectation by the following formula:

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega) P(\{\omega\}).$$

Law Of The Unconscious Statistician

- For a real function g and a discrete r.v. X, the function $Y = g(X)$ is also a discrete r.v.

Theorem (LOTUS)

For a real function g and a discrete r.v. X, we have

$$\mathbb{E}(g(X)) = \sum_{x \in Im(X)} g(x) P(X = x)$$

whenever the sum is defined.

Properties of $\mathbb{E}$

Theorem

Suppose X, Y are discrete r.v. and $a, b \in \mathbb{R}$.

1. *If $P(X \geq 0) = 1$ and $\mathbb{E}(X) = 0$, then $P(X = 0) = 1$.*
2. *If $\mathbb{E}(X) \geq 0$ then $P(X \geq 0) > 0$.*
3. $\mathbb{E}(a \cdot X + b) = a \cdot \mathbb{E}(X) + b$.
4. $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

Variance

Definition

Variance of a r.v. X is the number $\mathbb{E}((X - \mathbb{E}X)^2)$. It is denoted by $var(X)$.

Theorem

$$var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

Conditional expectation

Definition

Let X be a discrete r.v. and $P(B) > 0$. Conditional expectation of X given B is

$$\mathbb{E}(X \mid B) = \sum_{x \in Im(X)} x \cdot P(X = x \mid B),$$

whenever the sum is defined.

Law Of Total Expectation

Theorem

Suppose $B_1, B_2, \ldots$ *is a partition of* Ω *and* $A \in \mathcal{F}$. *Then*

$$\mathbb{E}(X) = \sum_i \mathbb{E}(X \mid B_i) P(B_i),$$

whenever the sum is defined. (Terms with $P(B_i) = 0$ *are counted as 0.)*

Law Of Total Expectation