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 $\boxed{\text{discrete}}$ random variable, if $Im(X)$ (range of X) is a countable set and if for every real x we have

$$X^{-1}(x) = \{\omega \in \Omega : X(\omega) = x\} \in \mathcal{F}.$$

$P(X \text{ assumes value } x) = P(X = x) =$ this is only defined if $\{X = x\} \in \mathcal{F}$

PMF

Definition

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

$Im\, X = \{1,2,3,4,5\}$

$p_X(1) = P(\{X=1\})$

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

$p_X(2)$

$\Omega = X^{-1}(1) \cup X^{-1}(2) \cup \dots \cup X^{-1}(5)$

$= \bigcup_{x \in Im X} X^{-1}(x)$

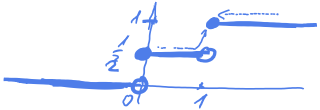
- $\sum_{x \in Im(X)} p_X(x) = ?$ 1
- $S := Im(X) \qquad Q(A) := \sum_{x \in A} p_X(x)$
  $(S, \mathcal{P}(S), Q)$ is a discrete probability space.

$1 = P(\Omega) = \sum P(X^{-1}(x))$

$= \sum P(X = x)$

- 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

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

$X = \begin{cases} 0 & \text{w. prob } \frac{1}{2} \\ 1 & \frac{1}{2} \end{cases}$

- F_X is a nondecreasing function → Proof $x < y \Rightarrow F_X(x) \le F_X(y)$?
- $\lim_{x \to -\infty} F_X(x) = 0$
- $\lim_{x \to +\infty} F_X(x) = 1$
- F_X is right-continuous

$\Leftrightarrow P(X \le x) \le P(X \le y)$

$A = \{\omega : X(\omega) \le x\}$ $\quad P(A) \le P(B)$

$B : \{\omega : X(\omega) \le y\}$ $\quad A \subseteq B$

$\omega \in A \Rightarrow X(\omega) \le x < y \Rightarrow \omega \in B$

Proof $A_n = \{X \le n\}$

$A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots$

$\bigcup_{n \ge 1} A_n = \Omega \Rightarrow P(\Omega) = \lim_{n \to \infty} P(A_n) = \lim_{n \to \infty} F_X(n)$

$\forall \varepsilon > 0 \ \exists n \in \mathbb{N} \ F_X(n) > 1 - \varepsilon \Rightarrow \forall x > n : F_X(x) > 1 - \varepsilon$, ≤ 1

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 $Alt(p)$.)

$X \rightsquigarrow A = X^{-1}(1)$

$X = I_A \leftsquigarrow A$

- Given $p \in [0, 1]$.
- $p_X(1) = p$ — $P(X=1)$
- $p_X(0) = 1 - p$ — $P(X=0)$
- $p_X(k) = 0$ 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_i = "i-th toss was a T"

- $X = \sum_{i=1}^{n} X_i$ 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 $k \in \{0, 1, \ldots, n\}$

$\| \quad P(X=k) =$ # of k-subsets of 1...n

P(k× tail)

P((n-k)× head)

k tails

H T H T T H H

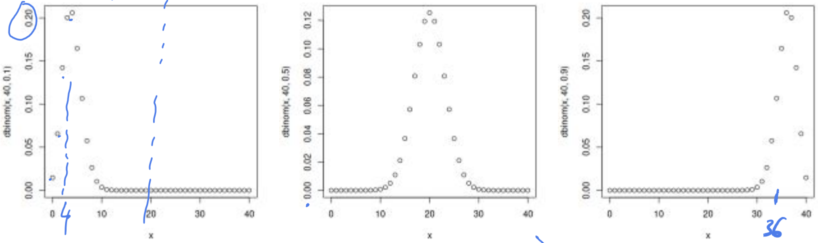
n tosses

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \overset{?}{=} 1$$

$$\| \quad (p + (1-p))^n = 1^n$$

Binomial distribution: PMF

Bin(40, 0.1) Bin(40, 0.9)



Bin(40, 1/2)

$$\binom{n}{k}\frac{1}{2^k}\cdot\frac{1}{2^{n-k}} = \frac{\binom{n}{k}}{2^n}$$

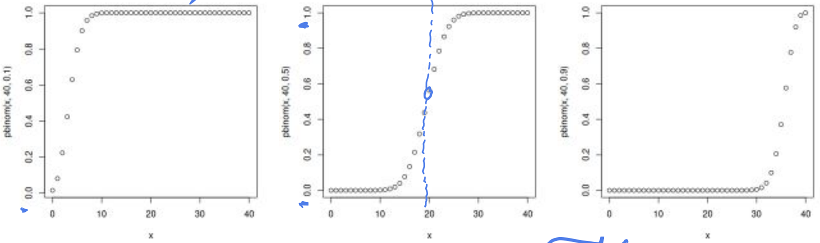
Generated by the following code in R

```
x <- 0:40              c(0,1,2,3,...,40)
plot(x,dbinom(x,40,0.1))
plot(x,dbinom(x,40,0.5))
plot(x,dbinom(x,40,0.9))
```

values vector   n   p

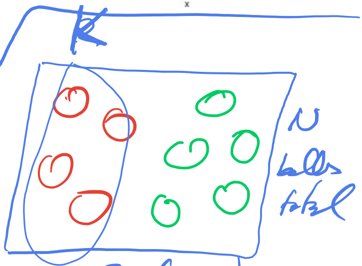
Binomial distribution: CDF

Bin(40, 0.1) Bin(40, 0.5) Bin(40, 0.9)



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))
```



Take n balls out of the box, with replacement

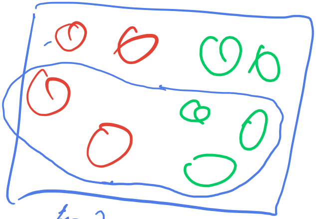
X: # of red

$X \sim Bin(n, p)$

$p = \frac{4}{9} = \frac{K}{N}$

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 $X \sim Hyper(N, K, n)$.
- $p_X(k) = P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$



$k=2$ $n=5$

set of k red balls

set of $n-k$ green balls

set of n balls out of N

$$\sum \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}} = 1$$

known from other classes

Poisson distribution

check $\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \overset{?}{=} 1$

$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \overset{?}{=} e^{\lambda}$

(True ... see calculus)

- We write $X \sim Pois(\lambda)$.

- Given real $\lambda > 0$.
- $p_X(k) = \frac{\lambda^k}{k!} e^{-\lambda}$ $k = 0, 1, 2, \ldots$ ≥ 0

- $Pois(\lambda)$ is a limit of $X_n \sim Bin(n, \lambda/n)$ fixed λ

Claim

- X describes, e.g., the number of emails we get in a day.

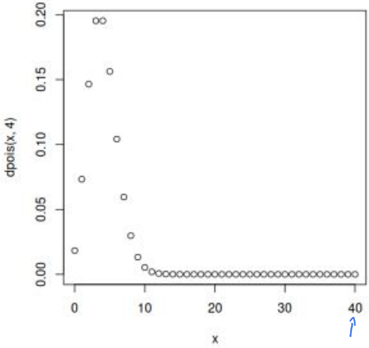
$\lim_{n\to\infty} P(X_n = k) = P(X = k)$ for every fixed k $\Rightarrow$ $\frac{n-k}{n} \to 1$

$$P(X_n = k) = \binom{n}{k}\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{n(n-1)\ldots(n-k+1)}{k!\, n^k} \lambda^k \cdot \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}$$

$\frac{n(n-1)\ldots(n-k+1)}{n^k} \to 1$, $\left(1 - \frac{\lambda}{n}\right)^n \to e^{-\lambda}$, $\left(1 - \frac{\lambda}{n}\right)^{-k} \to 1^k = 1$

$$\xrightarrow{n\to\infty} \frac{\lambda^k}{k!} \cdot 1 \cdot e^{-\lambda} \cdot 1 = P(X = k)$$

Poisson distribution: PMF



$\frac{\lambda^k}{k!} e^{-\lambda} > 0$

$\forall k \in \mathbb{N}$

It goes on forever

Generated by the following code in R

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

same 0:40

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

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

$\Longrightarrow$ that's why $Pois(\lambda)$ is good to model rare events

Geometric distribution

(independent)

Geom$_1$

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

p = prob. of a tail.

P(head (k-1) times)

P(tail at the final step)

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

"$P(X=k)$"

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

Geom$_0$

Overview

Discrete random variables

Examples of discrete r.v.'s

Expectation

Expectation

$0 \cdot P(X=0) + 1 \cdot P(X=1) + 100 \cdot P(X=100) + \dots$

$1 + \frac{1}{2} + \frac{1}{4} + \dots = 2$

$\sum_{n=1} (-1)^n$ undef.

$1+1+1+\dots = +\infty$

Definition

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

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

whenever the sum is defined.

people undf.

weight

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

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

how to

weighted average

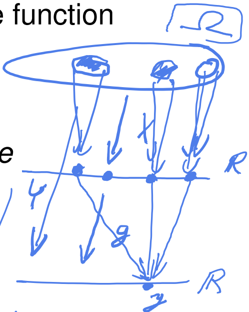
countable

$$(*) = \sum_{x \in Im(X)} \sum_{\substack{\omega \in \Omega \\ X(\omega) = x}} \underbrace{X(\omega)}_{= x} P(\{\omega\}) = \sum_{x \in Im X} \Big(x \cdot \sum_{\omega : X(\omega) = x} P(\{\omega\}) \Big) = \sum x \cdot P(\{\omega : X(\omega) = x\})$$

Law Of The Unconscious Statistician

$Y(\omega) = g(X(\omega))$

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



Theorem (LOTUS)

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

$$\underbrace{E(Y)}_{\text{def}} = \mathbb{E}(g(X)) = \sum_{x \in Im(X)} g(x) \underbrace{P(X = x)}_{\text{pmf of } X}$$

whenever the sum is defined.

Proof

$$E(Y) = \sum_{y \in Im\, Y} y \cdot \underline{P(Y = y)}$$

$$= \sum_{y \in Im\, Y} \Big(y \cdot \sum_{\substack{x \in Im\, X \\ g(x) = y}} P(X = x) \Big)$$

$$= \sum_{y} \sum_{x} g(x) P(X = x)$$

(1) Im Y is countable: $|Im\, Y| \le |Im\, X|$

(2) $\{\omega : Y(\omega) = y\} \in \mathcal{F}$

$$\{Y = y\} = \bigcup_{\substack{x \in Im\, X \\ g(x) = y}} \{\omega : X(\omega) = x\} \in \mathcal{F} \text{ as } X \text{ is a r.v.}$$

countable

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 $B_1, B_2, \dots$ 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