TDAB01 Probability and Statistics

Maryna Prus
IDA, Linköping University

Lecture 3: Families of Discrete Distributions

Overview

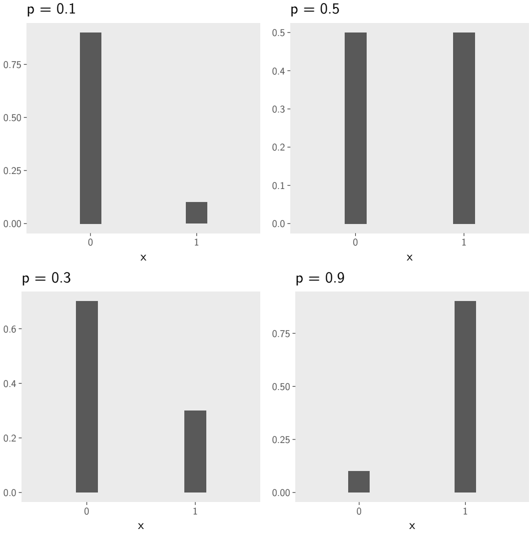
- **Bernoulli and Binomial distributions**
- **Geometric and Negative Binomial distributions**
- **Poisson distribution**

Bernoulli distribution

Definition. A random variable X with two possible values, 0 and 1, is called **Bernoulli variable**.
This random variable X is **Bernoulli distributed** and it holds that
$\boldsymbol{P}(X=1)=P(1)=p$ and $\boldsymbol{P}(X=0)=P(0)=q=1-p$.

- Notation: $X \sim Bernoulli(p)$ or $X \sim Be(p)$
- p - probability of success
- Any experiment with binary outcome is called Bernoulli trial
- By changing parameter p we obtain a variety (or family) of probability distributions
  A family of distributions is a variety of probability distributions that differ from each other by values of their parameters
- For R-codes for graphics in this lecture see `ManipDistributions.R`

Bernoulli distribution



Bernoulli distribution

- Pmf for $X \sim Bernoulli(p)$

$$P(x) = \begin{cases} q = 1 - p & \text{if x=0} \\ p & \text{if x=1} \end{cases}$$

- Expected value and variance of $X \sim Bernoulli(p)$

$$\mathbb{E}(X) = 0 \cdot q + 1 \cdot p = p$$
$$Var(X) = (0 - p)^2 \cdot q + (1 - p)^2 p = p - p^2 = p \cdot q$$

Binomial distribution

Definition. Let X be the number of successes in n **independent** Bernoulli trials with success probability p. Then X has **Binomial distribution** with parameters n and p and pmf

$$P(x) = \binom{n}{x} p^x q^{n-x}$$

for $x = 0, 1, 2, \ldots, n$.

- Notation: $X \sim Binomial(n, p)$, $X \sim Bin(n, p)$ or $X \sim Bi(n, p)$
- $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ - number of sequences of length n with exactly x successes
  $\binom{n}{x}$ is called **Binomial coefficient**
- Example: $n = 3$
  Sequences $(0, 1, 1)$, $(1, 0, 1)$ and $(1, 1, 0)$ lead to $x = 2$
  - Sequence $(0, 1, 1)$ has probability $q \cdot p \cdot p = p^2 q$
  - Sequence $(1, 0, 1)$ has probability $p \cdot q \cdot p = p^2 q$
  - Sequence $(1, 1, 0)$ has probability $p \cdot p \cdot q = p^2 q$

  Then $P(2) = 3 \cdot p^2 q$

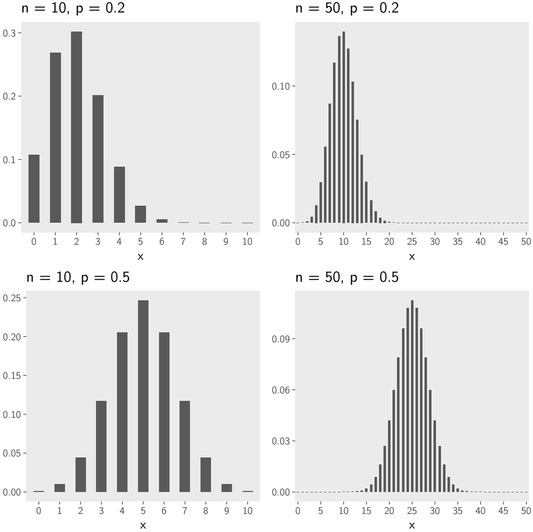
Binomial distribution

- Binomial distribution fits following data:
  - **discrete non-negative integers**
  - which can take all **integer values between** 0 **and** n
  - Suitable: How many students in class 5A can swim?
  - Not suitable: How many persons enter a shopping center on Saturday? (no natural upper limit); Length measurements (continuous)
- $X \sim Binomial(n, p)$ can be represented as sum of n independent Bernoulli variables $X_i \sim Bernoulli(p)$, $i = 1, \ldots, n$:

$$X = X_1 + X_2 + \cdots + X_n$$

  Note that success probability p has to be the same for both distributions
- Expectation and variance of $X \sim Binomial(n, p)$
  - $\mathbb{E}(X) = n \cdot p$
  - $Var(X) = n \cdot p \cdot q$
- Example: See Example 3.17 in textbook
- Values for cdf for all distributions in Appendix 3 in textbook

Binomial distribution



Geometric distribution

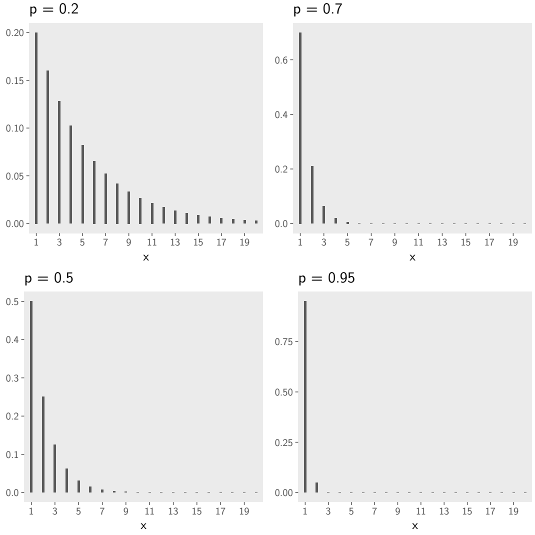
Definition. Let X be the number of **independent** Bernoulli trials needed to get the first success. Let p be the success probability. Then X is **Geometrically distributed** with pmf

$$P(x) = (1-p)^{x-1}p$$

for $x = 1, 2, \ldots$

- Notation: $X \sim Geometric(p)$ or $X \sim Geo(p)$
- Geometric distribution fits following data:
  - **discrete positive integers:** $1, 2, 3, \ldots$
  - which **do not have an upper limit** (in contrast to Binomial distribution)
- Expectation and variance of $X \sim Geo(p)$
  - $\mathbb{E}(X) = 1/p$
  - $Var(X) = \frac{1-p}{p^2}$
- Coin tossing (success="head"): $\mathbb{E}(X) = 2,\ Var(X) = 2$
- Rolling a die (success="6"): $\mathbb{E}(X) = 6,\ Var(X) = 30$

Geometric distribution



Negative Binomial distribution

Definition. Let X be the number of **independent** Bernoulli trials needed to get the k successes. Let p be the success probability. Then X has **Negative Binomial distribution** with pmf

$$P(x) = \binom{x-1}{k-1}(1-p)^{x-k}p^{k}$$

for $x = 1, 2, \dots$

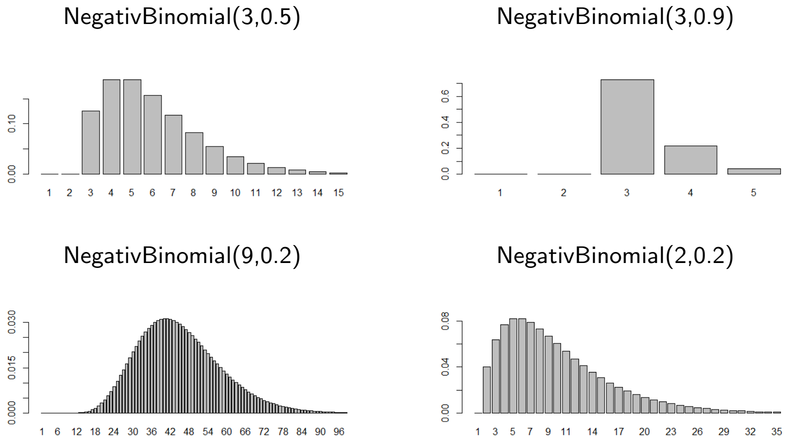
- Notation: $X \sim NegativBinomial(k, p)$, $X \sim NegBi(k, p)$
- $NegativBinomial(1, p) = Geo(p)$
- Negative Binomial distribution - opposite of Binomial distribution:
  - Binomial distribution - number of successes within n trials
  - Negative Binomial distribution - number of trials needed for k successes
- Example: See Example 3.21 in textbook

Negative Binomial distribution

- Negative Binomial distribution fits following data:
  - **discrete positive integers:** $1, 2, 3, \dots$
  - which **do not have an upper limit** (in contrast to Binomial distribution)
- Expected value and variance of $X \sim NegativBinomial(k, p)$
  - $\mathbb{E}(X) = k/p$
  - $Var(X) = \frac{k(1-p)}{p^2}$
- $X \sim NegativBinomial(k, p)$ can be represented as sum of k independent Geometrically distributed random variables $X_i \sim Geo(p)$, $i = 1, \dots, k$:

$$X = X_1 + X_2 + \dots + X_k$$

Negative Binomial distribution



Poisson distribution

Definition. **Poisson distributed** random variable with parameter λ has pmf

$$P(x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

for $x = 0, 1, 2, \dots$

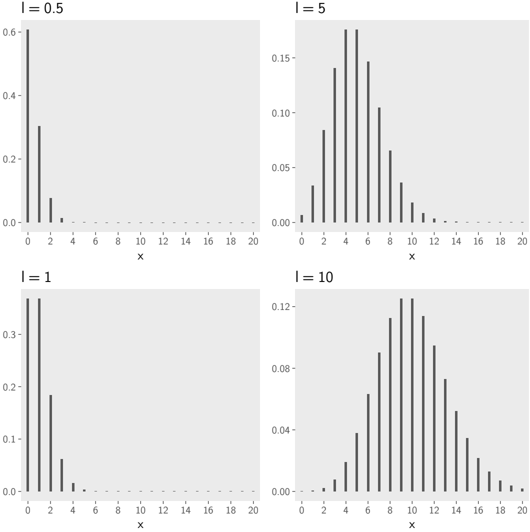
- Notation: $X \sim Po(\lambda)$
- Expected value and variance of $X \sim Po(\lambda)$
  - $\mathbb{E}(X) = \lambda$
  - $Var(X) = \lambda$
- Poisson distribution fits following data:
  - **Discrete non-negative integers:** $0, 1, 2, \dots$
  - which **do not have an upper limit**
  - expected value and variance are equal

Poisson distribution

- Number of unusual events within fixed period has Poisson distribution
- Examples:
  Number of detected bugs in a code
  The average number of major storms in your city is 2 per year. What is the probability that exactly 3 storms will hit your city next year?
    - X - number of storms next year
    - $\lambda = 2$, $X = 3$
    - $P(3) = \frac{e^{-2}2^3}{3!} \approx 0.18$

  Example 3.22 in textbook
- Poisson distribution with $\lambda = n \cdot p$ approximates Binomial distribution for large n ($n \geq 30$) and small p ($p \leq 0.05$)
  See `ManipDistributions.R`.

Poisson distribution



Thank you for your attention!