Discrete Mathematics and Probability
Week 8

Chris Heunen
Topics

- Bernoulli distribution: single trial
- Binomial distribution: many independent trials
- Poisson distribution: counting independent trials
- Geometric distribution: first success in independent trials
Bernoulli and Binomial distributions
Bernoulli distribution

Definition

Suppose that $n$ independent trials are performed, each succeeding with probability $p$. Let $X$ count the number of successes within the $n$ trials. Then $X$ has the Binomial distribution with parameters $n$ and $p$ or, in short, $X \sim \text{Binom}(n, p)$.

Special case $n = 1$ is called Bernoulli distribution with parameter $p$.

\[
\text{Bernoulli} = \text{indicator variable} \\
X = \begin{cases} 
1 & \text{if } E \text{ occurs} \\
0 & \text{if } \neg E \text{ occurs}
\end{cases}
\]
Bernoulli: mass function

**Proposition**

Let $X \sim \text{Binom}(n, p)$. Then $X = 0, 1, \ldots, n$, and its mass function is

$$p(i) = P(X = i) = \binom{n}{i} p^i (1 - p)^{n-i}, \quad i = 0, 1, \ldots, n.$$ 

In particular, the Bernoulli($p$) variable can take on values 0 or 1, with respective probabilities

$$p(0) = 1 - p, \quad p(1) = p.$$
Mass function

\[ p(i) \geq 0 \]

\[
\sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} = (p+(1-p))^n = 1
\]

**Newton's Binomial Theorem:**

\[
(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}
\]
Example

Screws are sold in packs of 10.
Each one is defective with probability 0.1.
If more than one screw is defective, you can return pack.
What percentage of packs is returned?

\[
X = \text{nr defective screws}
\]
\[
X \sim \text{Binom}(10, 0.1)
\]
\[
P(X \geq 2) = 1 - P(X = 0) - P(X = 1)
\]
\[
= 1 - \binom{10}{0} 0.1^0 0.9^0 - \binom{10}{1} 0.1^1 0.9^9 
\]
\[
\approx 26\
\]
Bernoulli: expectation, variance

**Proposition**

Let $X \sim \text{Binom}(n, p)$. Then:

$$EX = np, \quad \text{and} \quad \text{Var} \ X = np(1 - p)$$

**Proof:**

$$EX = \sum_i i \cdot P(i) = \sum_{i=0}^n i \cdot \binom{n}{i} p^i (1 - p)^{n-i}$$

$$= \sum_{i=0}^n \binom{n}{i} \left. \frac{d}{dt} t^i \right|_{t=1} \cdot p^i (1 - p)^{n-i}$$

$$= \frac{d}{dt} \left( \sum_{i=0}^n \binom{n}{i} (tp)^i (1 - p)^{n-i} \right) \bigg|_{t=1}$$

$$= \frac{d}{dt} (tp + 1-p)^n \bigg|_{t=1}$$

$$= n (tp + 1-p)^{n-1} \bigg|_{t=1}$$

$$= np$$

**Trick:**

$$i = \frac{d}{dt} t^i \bigg|_{t=1}$$
Proof

**Trick** \( i(i-1) = \frac{d^2}{dt^2} \frac{t^i}{t=1} \)

\[
E(x(x-1)) = \sum_{i=0}^{n} \binom{n}{i} i(i-1) p^i (1-p)^{n-i}
\]

\[
= \sum_{i=0}^{n} \binom{n}{i} \left. \frac{d^2}{dt^2} \frac{t^i}{t=1} \right|_{t=1} p^i (1-p)^{n-i}
\]

\[
= \left. \frac{d^2}{dt^2} \left( \sum_{i=0}^{n} \binom{n}{i} (tp)^i (1-p)^{n-i} \right) \right|_{t=1}
\]

\[
= \left. \frac{d^2}{dt^2} (tp + 1-p)^n \right|_{t=1}
\]

\[
= n(n-1) (tp + 1-p)^{n-2} p^2 \left. \right|_{t=1}
\]

\[
= n(n-1) p^2
\]

**Var** \( X 
\[
= E(X^2) - (E(X))^2
\]

\[
= E(X^2 - X) + E(X) - (E(X))^2
\]

\[
= n(n-1)p^2 + np - (np)^2
\]

\[
= np^2 - np^2 + np - np^2
\]

\[
= np(1-p)
\]
Poisson distribution
Poisson: mass function

The Poisson distribution is of central importance in Probability. Will later see relation to Binomial.

Definition

Fix a positive real number $\lambda$. The random variable $X$ is *Poisson distributed with parameter* $\lambda$, in short $X \sim \text{Poi}(\lambda)$, if it is non-negative integer valued, and its mass function is

$$p(i) = P(X = i) = e^{-\lambda} \frac{\lambda^i}{i!}, \quad i = 0, 1, 2, \ldots$$

Indeed a p.m.f. ✓

Fine, but why this distribution?
Poisson approximation of Binomial

**Proposition**

Fix \( \lambda > 0 \), and suppose that \( Y_n \sim \text{Binom}(n, p) \) with \( p = p(n) \) in such a way that \( n \cdot p \to \lambda \). Then the distribution of \( Y_n \) converges to \( \text{Poisson}(\lambda) \):

\[
\forall i \geq 0 \quad P(Y_n = i) \xrightarrow{n \to \infty} e^{-\lambda} \frac{\lambda^i}{i!}.
\]

Take \( Y \sim \text{Binom}(n, p) \) with large \( n \), small \( p \), such that \( np \approx \lambda \). Then \( Y \) is approximately \( \text{Poisson}(\lambda) \) distributed.
Proof

\[ P(Y_n = i) = \binom{n}{i} \rho^i (1-\rho)^{n-i} \]

\[ = \frac{1}{i!} \cdot \rho \cdot (n-1)\rho \cdots (n-i+1)\rho \cdot \frac{(1-\rho)^n}{(1-\rho)} \]

\[ \rightarrow \frac{1}{i!} \cdot \lambda^i \cdot e^{-\lambda} \]

\[ \lambda \rho \rightarrow \lambda \]

\[ (n-1)\rho \rightarrow \lambda \]

\[ \vdots \]

\[ (n-i+1)\rho \rightarrow \lambda \]

\[ (1-\rho)^n = \left(1 - \frac{1}{\rho}\right)^n \rightarrow e^{-\lambda} \]

\[ (1-\rho)^{n-i} \rightarrow \]
Poisson: expectation, variance

Proposition

For $X \sim \text{Poi} (\lambda)$, $EX = \text{Var} X = \lambda$.

(cf. $X \sim \text{Binom} (n, p)$  $EX = np$  $\text{Var} X = np(1-p)$ )

Proof:

$EX = \sum_{i=0}^{\infty} i P(i) = \sum_{i=0}^{\infty} i e^{-\lambda} \frac{\lambda^i}{i!}$

$= \lambda \sum_{i=0}^{\infty} i e^{-\lambda} \frac{\lambda^i}{(i-1)!} = \lambda \sum_{j=0}^{\infty} e^{-\lambda} \frac{\lambda^j}{j!} = \lambda$

$EX (X-1) = \sum_{i=0}^{\infty} i(i-1) P(i) = \sum_{i=0}^{\infty} i(i-1) e^{-\lambda} \frac{\lambda^i}{i!}$

$= \lambda^2 \sum_{i=0}^{\infty} e^{-\lambda} \frac{\lambda^{i-1}}{(i-1)!} = \lambda^2 \sum_{j=0}^{\infty} e^{-\lambda} \frac{\lambda^j}{j!} = \lambda^2$

$\text{Var} X = E(X^2) - (EX)^2$

$= E(X(X-1)) + EX - (EX)^2$

$= \lambda^2 + \lambda - \lambda^2$

$= \lambda$
Example

Poisson: many independent small probability events, summing up to "a few" in expectation

- nr of typos on a page of a book
- nr of citizens > 100 years of age in a city
- nr calls per hour in customer centre
- nr customers in post office today

book has average of \( \frac{1}{4} \) typos per page.

chance that next page has \( \geq 3 \) typos?

\( X \sim \text{Poisson}(\lambda) \)

\( \mathbb{E}X = \lambda = \frac{1}{4} \)

\[
\begin{align*}
p(X \geq 3) &= 1 - p(X \leq 2) \\
&= 1 - p(X = 0) - p(X = 1) - p(X = 2) \\
&= 1 - \frac{(\frac{1}{4})^0 \cdot e^{-\frac{1}{4}}}{0!} - \frac{(\frac{1}{4})^1 \cdot e^{-\frac{1}{4}}}{1!} - \frac{(\frac{1}{4})^2 \cdot e^{-\frac{1}{4}}}{2!} \\
&\approx 1.47%. 
\end{align*}
\]
Examples

Defective screws.

Let $X \sim \text{Binom}(10, 0.1)$

well approximated by $\text{Poisson}(1)$ as $\lambda = 1 = 10 \cdot 0.1 = np$

\[
P(X \geq 2) = 1 - P(X = 0) - P(X = 1)
\]

\[
= 1 - e^{-\lambda} \cdot \frac{\lambda^0}{0!} - e^{-\lambda} \cdot \frac{\lambda^1}{1!}
\]

\[
\approx 1 - e^{-1} \cdot \frac{1^0}{0!} - e^{-1} \cdot \frac{1^1}{1!} \approx 26\%
\]
Geometric distribution
Geometric: mass function

Again independent trials, but now ask: when is the first success?

Definition
Suppose that independent trials, each succeeding with probability \( p \), are repeated until the first success. The total number \( X \) of trials made has the Geometric\((p)\) distribution (in short, \( X \sim \text{Geom}(p) \)).

Proposition
\( X \) can take on positive integers, with probabilities
\[
p(i) = (1 - p)^{i-1} \cdot p, \quad i = 1, 2, \ldots
\]

Proof:
\[
p(i) > 0 \quad \forall i
\]
\[
\sum_{i=1}^{\infty} p(i) = \sum_{i=1}^{\infty} (1-p)^{i-1} \cdot p = \frac{p}{1-(1-p)} = 1
\]
Corollary

The Geometric random variable is (discrete) memoryless:

\[ P\{X \geq n + k \mid X > n\} = P\{X \geq k\} \]

for every \( k \geq 1, \ n \geq 0. \)
Geometric: expectation, variance

**Proposition**

For a Geometric$(p)$ random variable $X$:

$$\text{EX} = \frac{1}{p}, \quad \text{Var} \, X = \frac{1 - p}{p^2}$$

**Proof:**

$$\text{EX} = \sum_{i=1}^{\infty} i \cdot (1-p)^{i-1} \cdot p = \sum_{i=0}^{\infty} (1-p)^{i+1} \cdot p$$

$$= \sum_{i=0}^{\infty} \frac{d}{dt} \left( \frac{1}{1-(1-p)t} \right)_{t=1} \cdot (1-p)^{i+1} \cdot p = \frac{d}{dt} \left( \sum_{i=0}^{\infty} \frac{1}{1-(1-p)t} \cdot (1-p)^{i+1} \cdot p \right)_{t=1}$$

$$= \frac{p}{1-p} \cdot \frac{1}{1-(1-p)} \left|_{t=1} \right.$$  

$$= \frac{p}{1-p} \cdot \frac{1}{1-(1-p)} = \frac{1}{p}$$
Example

First 3 appearing on fair die takes $X \sim \text{Geom}\left(\frac{1}{6}\right)$ rolls.

$\text{E} X = \frac{1}{\frac{1}{6}} = 6$

$\text{SD} X = \sqrt{\text{Var} X} = \sqrt{\frac{1-\frac{1}{6}}{\left(\frac{1}{6}\right)^2}} = \sqrt{30} \approx 5.48$

Chance that first 3 comes on $7^{th}$ roll is

$p(7) = p(X = 7) = \left(1-\frac{1}{6}\right)^6 \cdot \frac{1}{6} = 5.6\%

or later;

$p(X \geq 7) = \left(1-\frac{1}{6}\right)^6 \approx 33.5\%$
Summary

- Bernoulli distribution: single trial
- Binomial distribution: many independent trials
- Poisson distribution: counting independent trials
- Geometric distribution: first success in independent trials