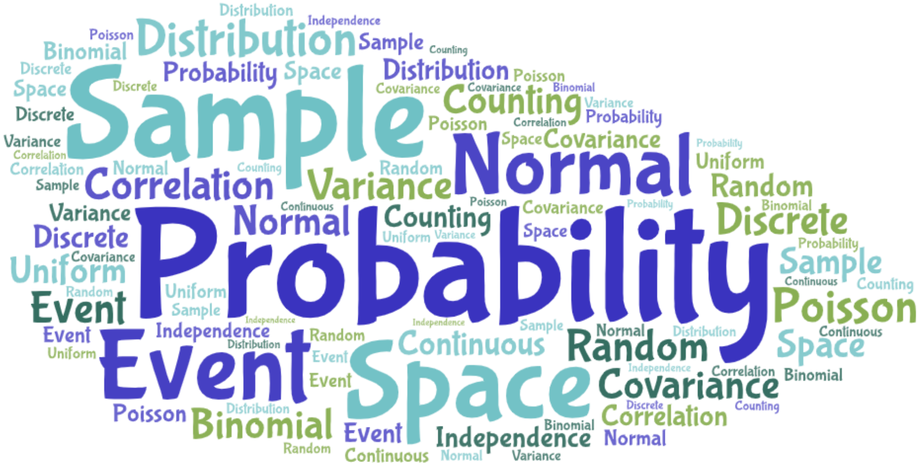
Discrete Mathematics and Probability

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, $X \sim Binom(n, p)$.

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

Bernoulli = indicator variable

$$X = \begin{cases} 1 & \text{if } E \text{ occurs} \\ 0 & \text{if } E^c \text{ occurs} \end{cases}$$

Bernoulli: mass function

Proposition

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

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

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

$$\mathfrak{p}(0) = 1 - p, \qquad \mathfrak{p}(1) = p.$$

Mass function

$p(i) \geqslant 0$ ✓

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

Newton's Binomial Theorem:

$$(x+y)^{n} = \sum_{k=0}^{n} \binom{n}{k} \cdot x^{k} \cdot y^{n-k}$$

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?

$X =$ nr defective screws

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

$$P(X \geq 2) = 1 - P(X=0) - P(X=1)$$
$$= 1 - \binom{10}{0} 0.1^0 \, 0.9^{10} - \binom{10}{1} 0.1^1 \, 0.9^9 \approx 26\%$$

Bernoulli: expectation, variance

Proposition

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

$$\mathbf{E}X = np, \qquad \textit{and} \qquad \mathbf{Var}\, X = np(1-p)$$

Proof:

$$\mathbb{E}X = \sum_i i \cdot p(i) = \sum_{i=0}^{n} i \cdot \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i}$$

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

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

Newton

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

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

$$= np$$

Trick: $i = \frac{d}{dt} t^i \Big|_{t=1}$

Proof

Trick $\quad i(i-1) = \frac{d^2}{dt^2} t^i \Big|_{t=1}$

$$
\begin{aligned}
\mathbb{E}(X(X-1)) &= \sum_{i=0}^{n} \binom{n}{i} i(i-1) p^i (1-p)^{n-i} \\
&= \sum_{i=0}^{n} \binom{n}{i} \frac{d^2}{dt^2} t^i \Big|_{t=1} p^i (1-p)^{n-i} \\
&= \frac{d^2}{dt^2} \left(\sum_{i=0}^{n} \binom{n}{i} (tp)^i (1-p)^{n-i} \right) \Big|_{t=1} \\
&= \frac{d^2}{dt^2} (tp + 1-p)^n \Big|_{t=1} \\
&= n(n-1)(tp+1-p)^{n-2} p^2 \Big|_{t=1} \\
&= n(n-1)p^2
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}\, X &= \mathbb{E}X^2 - (\mathbb{E}X)^2 \\
&= \mathbb{E}(X^2 - X) + \mathbb{E}X - (\mathbb{E}X)^2 \\
&= n(n-1)p^2 + np - (np)^2 \\
&= \cancel{n^2p^2} - np^2 + np - \cancel{n^2p^2} \\
&= np(1-p)
\end{aligned}
$$

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 λ. The random variable X is *Poisson distributed with parameter* λ, in short $X \sim Poi(\lambda)$, if it is non-negative integer valued, and its mass function is

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

Indeed a p.m.f. ✓

Fine, but why this distribution?

Poisson approximation of Binomial

Proposition

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

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

Take $Y \sim Binom(n,p)$ with large n, small p, such that $np \approx \lambda$

Then Y is approximately $Poisson(\lambda)$ distributed.

Proof

$$P(Y_n = i) = \binom{n}{i} p^i (1-p)^{n-i}$$

$$= \frac{1}{i!} \cdot np \cdot (n-1)p \cdots (n-i+1)p \cdot \frac{(1-p)^n}{(1-p)^i}$$

$$\xrightarrow[n\to\infty]{} \frac{1}{i!} \cdot \lambda^i \cdot e^{-\lambda}$$

$$np \to \lambda$$
$$(n-1)p \to \lambda$$
$$\vdots$$
$$(n-i+1)p \to \lambda$$
$$(1-p)^n = \left(1 - \frac{1}{1/p}\right)^n \to e^{-\lambda}$$
$$(1-p)^i \to 1$$

Poisson: expectation, variance

Proposition

For $X \sim Poi(\lambda)$, $\mathbf{E}X = \mathbf{Var}\, X = \lambda$.

(cf. $X \sim Binom(n,p)$ $\quad \mathbb{E}X = np \quad \mathbb{V}ar\, X = np(1-p)$)

Proof: $\mathbb{E}X = \sum_{i=0}^{\infty} i\,p(i) = \sum_{i=1}^{\infty} i \cdot e^{-\lambda} \cdot \frac{\lambda^i}{i!}$

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

$$\mathbb{E}(X(X-1)) = \sum_{i=0}^{\infty} i(i-1)p(i) = \sum_{i=2}^{\infty} i(i-1)e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$

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

$$\begin{aligned} \mathbb{V}ar\, X &= \mathbb{E}X^2 - (\mathbb{E}X)^2 \\ &= \mathbb{E}(X(X-1)) + \mathbb{E}X - (\mathbb{E}X)^2 \\ &= \lambda^2 + \lambda - \lambda^2 \\ &= \lambda \end{aligned}$$

Example

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

e.g.: - 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 $1/2$ typos per page.
chance that next page has ≥ 3 typos?

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

$\mathbb{E}X = \lambda = 1/2$

$$P(X \geq 3) = 1 - P(X \leq 2)$$
$$= 1 - P(X=0) - P(X=1) - P(X=2)$$
$$= 1 - \frac{(1/2)^0}{0!} \cdot e^{-1/2} - \frac{(1/2)^1}{1!} \cdot e^{-1/2} - \frac{(1/2)^2}{2!} \cdot e^{-1/2} \approx 1.4\%$$

Examples

Defective screws.

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

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

$$P(X \geq 2) = 1 - P(X=0) - P(X=1)$$

$$\simeq 1 - e^{-1} \cdot \frac{1^0}{0!} - e^{-1} \cdot \frac{1^1}{1!} \simeq 26\%$$

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, $X \sim Geom(p)$).

Proposition

X can take on positive integers, with probabilities $\mathfrak{p}(i) = (1-p)^{i-1} \cdot p$, $i = 1, 2, \ldots$.

Proof: $p(i) \geq 0$ ✓

$$\sum_{i=1}^{\infty} p(i) = \sum_{i=1}^{\infty} (1-p)^{i-1} \cdot p = \frac{p}{1-(1-p)} = 1$$

Geometric: mass function

Note: if $k \geq 1$ then $P(X \geq k) = (1-p)^{k-1}$: at least $k-1$ failures

Corollary

The Geometric random variable is (discrete) memoryless:

$$\mathbf{P}\{X \geq n+k \,|\, X > n\} = \mathbf{P}\{X \geq k\}$$

for every $k \geq 1$, $n \geq 0$.

Geometric: expectation, variance

Proposition

For a Geometric(p) random variable X:

$$\mathbf{E}X = \frac{1}{p} \qquad\qquad \mathbf{Var}\,X = \frac{1-p}{p^2}$$

Proof: $\mathbb{E}X = \sum_{i=1}^{\infty} i \cdot (1-p)^{i-1} \cdot p = \sum_{i=0}^{\infty} i \cdot (1-p)^{i-1} \cdot p$

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

$$= \frac{p}{1-p} \cdot \frac{d}{dt} \frac{1}{1-(1-p)t}\Big|_{t=1}$$

$$= \frac{p}{1-p} \cdot \frac{1-p}{(1-(1-p))^2} = \frac{1}{p}$$

Example

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

$$\mathbb{E}X = \frac{1}{1/6} = 6$$

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

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

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

or later:

$$P(X \geq 7) = \left(1 - \frac{1}{6}\right)^6 \simeq 33.5\%$$

Summary

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