Introduction to Probability

Lecture 4: More discrete distributions – Poisson, Geometric, Negative Binomial, Hypergeometric

Mateja Jamnik, Thomas Sauerwald

University of Cambridge, Department of Computer Science and Technology
email: {mateja.jamnik,thomas.sauerwald}@cl.cam.ac.uk

UNIVERSITY OF CAMBRIDGE

Outline

Poisson discrete random variable

Geometric discrete random variable

Negative binomial discrete random variable

Hypergeometric discrete random variable

Preliminaries:

The natural exponent *e*

e is a mathematical constant AKA the Euler number. *e* is very important for exponential functions. Here are some important identities:

$$e \approx 2.71828$$

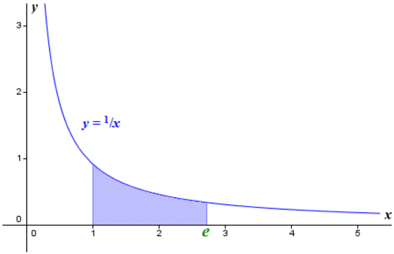
$$e = \sum_{n=0}^{\infty} \frac{1}{n!}$$

$$e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$$

$$e^{-\lambda} = \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n$$

$$e^r = \lim_{n\to\infty} \left(1 + \frac{r}{n}\right)^n$$

$$e^r = \sum_{n=0}^{\infty} \frac{r^n}{n!}$$



Binomial RV example: large n, small p

We are trying to predict footfall in a store. We know, based on previous data, that on average 8 people enter the store per hour. What is the probability of k people entering the store in the next 1 hour?

What if 2 people enter in the same minute?

1. Break an hour into **minutes**.
   - At each **minute**, independent Bernoulli trial with 1 for a person entering the store and 0 for nobody entering the store.
   - X is a Binomial RV: # people entering in an hour, so $\mathbf{E}[X] = np = \lambda = 8$.
   - $X \sim Bin(n = 60, p = \frac{\lambda}{n})$, so $\mathbf{P}[X{=}k] = \binom{n}{k} p^k (1-p)^{n-k} = \binom{60}{k} \left(\frac{8}{60}\right)^k \left(1 - \frac{8}{60}\right)^{n-k}$

What if 2 people enter in the same millisecond?

2. Break an hour into **milliseconds**.
   - At each **millisecond**, independent Bernoulli trial: 1 for enter, 0 for not enter.
   - X is a Binomial RV: # people entering in an hour, so $\mathbf{E}[X] = np = \lambda = 8$.
   - $X \sim Bin(n = 3600000, p = \frac{\lambda}{n})$, so $\mathbf{P}[X{=}k] = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$

3. Break an hour into **infinitely small units**.
   - At each **unit**, independent Bernoulli trial: 1 for enter, 0 for not enter.
   - X is a Binomial RV: # people entering in an hour, so $\mathbf{E}[X] = np = \lambda = 8$.
   - $X \sim Bin(n, p = \frac{\lambda}{n})$, thus $\mathbf{P}[\mathbf{X{=}k}] = \lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$

Computing Binomial in the limit

$$\mathbf{P}[X = k] = \lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \underset{expand}{=} \lim_{n\to\infty} \frac{n!}{k!(n-k)!} \frac{\lambda^k}{n^k} \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k}$$

$$\underset{rearrange}{=} \lim_{n\to\infty} \frac{n!}{n^k(n-k)!} \frac{\lambda^k}{k!} \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k} \underset{def\ of\ e}{=} \lim_{n\to\infty} \frac{n!}{n^k(n-k)!} \frac{\lambda^k}{k!} \frac{e^{-\lambda}}{\left(1-\frac{\lambda}{n}\right)^k}$$

$$\underset{expand}{=} \lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+1)}{n^k} \frac{\lambda^k}{k!} \frac{e^{-\lambda}}{\left(1-\frac{\lambda}{n}\right)^k}$$

as $n \to \infty$

$$\frac{n(n-1)\cdots(n-k+1)}{n^k} \approx \frac{n^k}{n^k} = 1$$

$$\left(1 - \frac{\lambda}{n}\right)^k \approx 1^k = 1$$

$$\left(1 - \frac{\lambda}{n}\right)^n \approx e^{-\lambda}$$ because $e^{-\lambda} = \lim_{n\to\infty}\left(1 - \frac{\lambda}{n}\right)^n$ thus we have

$$\mathbf{P}[X = k] = (1)\frac{\lambda^k}{k!}\frac{e^{-\lambda}}{1} = \frac{\lambda^k}{k!}e^{-\lambda}$$

Therefore, in our store footfall example: the probability of k people entering the store in the next 1 hour is:

$$\mathbf{P}[X=k] = \lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} = \frac{\lambda^k}{k!}e^{-\lambda}$$

Poisson

Poisson discrete random variable

A Poisson RV X approximates Binomial where n is large, p is small, and $\lambda = np$ is "moderate". Thus we no longer need to know n and p, we only need to provide **rate** λ. X is the number of successes over the duration of the experiment.

$$\mathbf{X}\sim\mathbf{Pois}(\lambda)$$

$$\begin{aligned} \text{Range:} &\quad \{0, 1, 2, \ldots\} \\ \text{PMF:} &\quad \mathbf{P}[\, X = k \,] = \frac{\lambda^k}{k!} e^{-\lambda} \\ \text{Expectation:} &\quad \mathbf{E}[\, X \,] = \lambda \\ \text{Variance:} &\quad \mathbf{V}[\, X \,] = \lambda \end{aligned}$$

Examples: # earthquakes in a given year, # goals scored during a 90 minute football game, # misprints per page in a book, # emails per day.

Key idea: Divide time into a **large number** of small increments. Assume that during each increment, there is some **small probability** of the event happening (independent of other increments).

Earthquake example

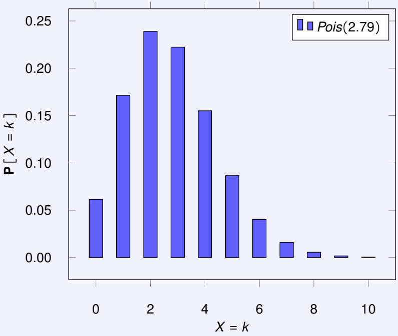
Example

Suppose there are an average of 2.79 major earthquakes in the world each year. What is the probability of getting 3 major earthquakes next year?

Answer

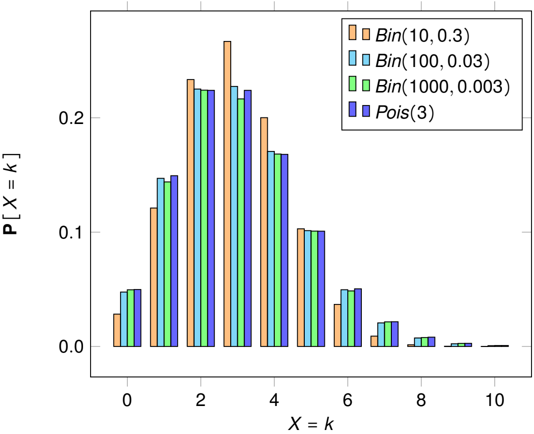
Define RVs: $\lambda = 2.79, k = 3, X \sim Pois(2.79)$

Solve: $\mathbf{P}[X = k] = \frac{\lambda^k}{k!}e^{-\lambda} = \mathbf{P}[X = 3] = \frac{2.79^3}{3!}e^{-2.79} \approx 0.23$



Poisson paradigm

- Poisson approximates Binomial when n is large, p is small, and $\lambda = np$ is "moderate".
- Different interpretations of "moderate". Commonly accepted ranges are:
  - $n > 20$ and $p < 0.05$
  - $n > 100$ and $p < 0.1$
- Poisson is Binomial in the limit: $\lambda = np$ where $n \to \infty, p \to 0$.



Poisson expectation

$$\text{PMF: } = k \in \{0, 1, 2, \ldots \infty\}; \mathbf{P}[\, X = k \,] = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$\mathbf{E}[\, X \,] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} =$$

$$= \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} \text{ (let } i = k-1\text{)}$$

$$= \lambda e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

Poisson variance

$$\mathbf{E}\left[X^2\right] = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \lambda \sum_{k=1}^{\infty} k \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} \text{ (let } i = k-1\text{)}$$

$$= \lambda \sum_{i=0}^{\infty} (i+1) \frac{\lambda^i}{i!} e^{-\lambda} = \lambda \left(\underbrace{\sum_{i=0}^{\infty} i \frac{\lambda^i}{i!} e^{-\lambda}}_{\textit{same as before}} + \underbrace{\sum_{i=0}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda}}_{\textit{sum of PMFs}=1} \right) =$$

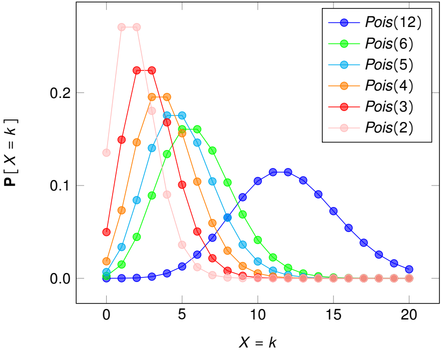
$$= \lambda(\lambda + 1) \qquad \text{thus}$$

$$\mathbf{V}[X] = \mathbf{E}\left[X^2\right] - (\mathbf{E}[X])^2 = \lambda(\lambda+1) - \lambda^2 = \lambda$$

$$\mathbf{E}\left[X^k\right] = \lambda \mathbf{E}\left[(X+1)^{k-1}\right]$$

Bernoulli, Poisson, and random processes

- A Poisson process is a model for a series of discrete events where the average time between events is known, but the exact timing of events is random.
  - The arrival of an event is independent of the event before (waiting time between events is memoryless).
  - The average rate (events per time period) is constant.
  - Two events cannot occur at the same time: each sub-interval of a Poisson process is a Bernoulli trial that is either a success or a failure.
- Example: your website goes down on average twice per 60 days; calling a help centre; movements of stock price...



Outline

Geometric

Geometric discrete random variable

X is a geometric RV if X is a number of independent Bernoulli trials until the first success, and p is the probability of success on each Bernoulli trial.

$$\mathbf{X \sim Geo(p)}$$

$$\begin{aligned} \text{Range:} &\quad \{1, 2, \ldots\} \\ \text{PMF:} &\quad \mathbf{P}[\,X = n\,] = (1-p)^{n-1} p \\ \text{Expectation:} &\quad \mathbf{E}[\,X\,] = \frac{1}{p} \\ \text{Variance:} &\quad \mathbf{V}[\,X\,] = \frac{1-p}{p^2} \end{aligned}$$

Examples: tossing a coin ($\mathbf{P}[\,head\,] = p$) until first heads appears, generating bits with $\mathbf{P}[\,bit = 1\,] = p$ until first 1 is generated.

Geometric RV properties

PMF (E_i is the event that the i-th trial succeeds):

$$\begin{aligned}\mathbf{P}[X = n] &= \mathbf{P}[E_1^c E_2^c \dots E_{n-1}^c E_n] = \\ &= \mathbf{P}[E_1^c]\mathbf{P}[E_2^c]\dots\mathbf{P}[E_{n-1}^c]\mathbf{P}[E_n] = \\ &= (1-p)^{n-1}p\end{aligned}$$

CDF ($\mathbf{P}[X > n]$ is the probability that at least the first n trials fail):

$$\begin{aligned}\mathbf{P}[X \le n] &= 1 - \mathbf{P}[X > n] = \\ &= 1 - \mathbf{P}[E_1^c E_2^c \dots E_n^c] = \\ &= 1 - \mathbf{P}[E_1^c]\mathbf{P}[E_2^c]\dots\mathbf{P}[E_n^c] = \\ &= 1 - (1-p)^n\end{aligned}$$

Die example

Example

You roll a fair 6-sided die until it comes up with # 6. What is the probability that it will take 3 rolls?

Answer

Let X be a RV for # of rolls. Probability for any # on die is $\frac{1}{6}$.

Define RVs: $X \sim Geo(\frac{1}{6})$, want $\mathbf{P}[X = 3]$.

Solve:
$$\mathbf{P}[X = 3] = (1-p)^{n-1} p \text{ where } n = 3, p = \frac{1}{6}$$
$$= \left(1 - \frac{1}{6}\right)^{3-1} \frac{1}{6} = \left(\frac{5}{6}\right)^{2} \frac{1}{6} = \frac{25}{216}$$

$$\mathbf{P}[X = 3] = \mathbf{P}[\,not\ 6, not\ 6, 6\,] = \frac{5}{6}\frac{5}{6}\frac{1}{6} = \frac{25}{216}$$

$$\mathbf{P}[X = n] = \mathbf{P}[\,not\ 6 (n-1) times, 6\,] = \left(\frac{5}{6}\right)^{n-1}\left(\frac{1}{6}\right)$$

Outline

Negative binomial

Negative binomial discrete random variable

X is a negative binomial RV if X is the number of independent Bernoulli trials until r successes and p is the probability of success on each trial.

$$\mathbf{X \sim NegBin(r, p)}$$

Range: $\{r, r+1, \ldots\}$

PMF: $\mathbf{P}[X = n] = \binom{n-1}{r-1}(1-p)^{n-r}p^r$

Expectation: $\mathbf{E}[X] = \frac{r}{p}$

Variance: $\mathbf{V}[X] = \frac{r(1-p)}{p^2}$

Examples: tossing a coin until r-th heads appears, generating bits until the first r 1's are generated.

Note: $Geo(p) = NegBin(1, p)$.

NegBin example

Example (not real life!)

A PhD student is expected to publish 2 papers to graduate. A conference accepts each paper randomly and independently with probability $p = 0.25$. On average, how many papers will the student need to submit to a conference in order to graduate?

Answer

Let X be # submissions required to get 2 acceptances. Thus $X \sim NegBin(r = 2, p = 0.25)$. So,

$$\mathbf{E}[X] = \frac{r}{p} = \frac{2}{0.25} = 8$$

Adding NegBin example

Example

Let $X \sim NegBin(m, p)$ and $Y \sim NegBin(n, p)$ be two independent RVs. Define a new RV as $Z = X + Y$. Find PMF of Z.

Answer

- Need to show that $Z \sim NegBin(m + n, p)$.
- Consider the sequence of independent events tossing a coin with $\mathbf{P}[\, heads \,] = p$.
- Let X be a RV for # of coin tosses until m heads are observed. Thus $X \sim NegBin(m, p)$.
- Now, continue to toss a coin after m heads are observed, until n more heads are observed. Thus, for this part of the sequence, $Y \sim NegBin(n, p)$.
- Looking at it from the beginning we tossed independently the coin until we observed $m + n$ heads, thus $Z = X + Y$ and thus $Z \sim NegBin(m + n, p)$.
- Note: if $X_1, X_2, \ldots X_m$ are m independent $Geo(p)$ RVs, then the RV $X = X_1 + X_2 + \cdots + X_m$ has $NegBin(m, p)$ distribution.

Outline

Poisson discrete random variable

Geometric discrete random variable

Negative binomial discrete random variable

Hypergeometric discrete random variable

Hypergeometric

Hypergeometric discrete random variable

X is a hypergeometric RV that samples n objects, without replacement, with i successes (random draw for which the object drawn has a specified feature), from a finite population of size N that contains exactly m objects with that feature.

$$\mathbf{X} \sim \mathbf{Hyp}(\mathbf{N}, \mathbf{n}, \mathbf{m})$$

Range: $\{0, 1, \ldots, n\}$

PMF: $\mathbf{P}[\, X = i \,] = \dfrac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}}$

Expectation: $\mathbf{E}[\, X \,] = n\dfrac{m}{N}$

Variance: $\mathbf{V}[\, X \,] = n\dfrac{m}{N}\left(1 - \dfrac{m}{N}\right)\left(1 - \dfrac{n-1}{N-1}\right)$

Example: an urn has N balls of which m are white and $N - m$ are black; we take a random sample **without replacement** of size n and measure X: # of white balls in the sample.

Survey sampling

Example

A street has 40 houses of which 5 houses are inhabited by families with an income below the poverty line. In a survey, 7 houses are sampled at random from this street. What is the probability that: (a) none of the 5 families with income below poverty line are sampled? (b) 4 of them are sampled? (c) no more than 2 are sampled? (d) at least 3 are sampled?

Answer

Let X: # of families sampled which are below the poverty line.

$X \sim Hyp(N = 40, n = 7, m = 5)$.

(a) $\mathbf{P}[\,X = 0\,] = \dfrac{\binom{5}{0}\binom{40-5}{7-0}}{\binom{40}{7}} = \dfrac{\binom{35}{7}}{\binom{40}{7}} \approx 0.36$

(b) $\mathbf{P}[\,X = 4\,] = \dfrac{\binom{5}{4}\binom{40-5}{7-4}}{\binom{40}{7}}$

(c) $\mathbf{P}[\,X \le 2\,] = \mathbf{P}[\,X = 0\,] + \mathbf{P}[\,X = 1\,] + \mathbf{P}[\,X = 2\,]$

(d) $\mathbf{P}[\,X \ge 3\,] = 1 - \mathbf{P}[\,X \le 2\,]$

Summary of discrete RV

	$Ber(p)$	$Bin(n,p)$	$Pois(\lambda)$	$Geo(p)$	$NegBin(r,p)$	$Hyp(N,n,m)$
PMF	$\mathbf{P}[X=1]=p$	$\mathbf{P}[X=k]=\binom{n}{k}p^k(1-p)^{n-k}$	$\mathbf{P}[X=k]=\frac{\lambda^k}{k!}e^{-\lambda}$	$\mathbf{P}[X=n]=(1-p)^{n-1}p$	$\mathbf{P}[X=n]=\binom{n-1}{r-1}(1-p)^{n-r}p^r$	$\mathbf{P}[X=i]=\frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}}$
$\mathbf{E}[X]$	p	np	λ	$\frac{1}{p}$	$\frac{r}{p}$	$n\frac{m}{N}$
$\mathbf{V}[X]$	$p(1-p)$	$np(1-p)$	λ	$\frac{1-p}{p^2}$	$\frac{r(1-p)}{p^2}$	$n\frac{m}{N}\left(1-\frac{m}{N}\right)\left(1-\frac{n-1}{N-1}\right)$
Descr.	1 experiment with prob p of success	n independent trials with prob p of success	# successes over experiment duration, $\lambda = np$ rate of success	# independent trials until first success	# independent trials until r successes	# successes of drawing item with a feature (without replacement) in a sample of size n from a population of size N with m items with the feature