Today: Poisson distribution.
Expectation
Next: ASV 3.4

Week 5:

- homework 3 (due Wednesday October 26)
- Midterm 1 regrades open until October 30
Rare events. Poisson distribution

Let \( \lambda > 0 \) and let \( X \) be a r.v. taking values in \( \{0, 1, 2, \ldots\} \). 
\( X \) has Poisson distribution with parameter \( \lambda \) if
\[
P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{for } k \in \{0, 1, 2, \ldots\}
\]
We write \( X \sim \text{Pois}(\lambda) \).

Poisson distribution describes the probability that a "rare" event occurs \( k \) times after repeating the experiment (independent trials) "many" times.

Is this a probability distribution?

\[
P(X = k) \geq 0,
\]
\( \lambda \) gives the "expected number" of occurrences.
Rare events. Poisson distribution

Let $X$ be the number of successes in $n$ independent trials with success probability $\frac{\lambda}{n}$, $\lambda > 0$.

Then $P(X = k) =$

What happens if $(k \in \{0, 1, 2, \ldots, y \text{ is fixed}\})$?
**Poisson distribution. Example**

Observation: between 1875 and 1894 (20 years) in 14 units of Prussian army there were 196 deaths from horse kicks, distributed in the following way:

<table>
<thead>
<tr>
<th># deaths per unit per year, ( k )</th>
<th># unit-years with ( k ) deaths</th>
<th>empirical probability</th>
<th>( P(X=k) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>144</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5+</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>total</strong></td>
<td></td>
<td>280</td>
</tr>
</tbody>
</table>

Let is "expected number" of death per unit
Poisson distribution. Example

A 100 year storm is a storm magnitude expected to occur in any given year with probability $\frac{1}{100}$. Over the course of a century, how likely is it to see at least 4 100 year storms?
Summary

Independent trials: the most important (discrete) probability distributions are:

- **Ber** \( (p) \): \( P(X=1) = p, \ P(X=0) = 1-p \), \( 0 \leq p \leq 1 \)
  (single trial with success probability \( p \))

- **Bin** \( (n, p) \): \( P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k} \), \( 0 \leq k \leq n \)
  (number of successes in \( n \) independent trials with rate \( p \))

- **Geom** \( (p) \): \( P(N = k) = (1-p)^{k-1} p \), \( k = 1, 2, 3, ... \)
  (first successful trial in repeated independent trials with rate \( p \))

- **Poisson** \( (\lambda) \): \( P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \), \( k = 0, 1, 2, ... \) \( \lambda > 0 \)
  (approimates \( Bin(n, \frac{\lambda}{n}) \), number of rare events in many trials)
Expectation

Example: Toss a fair coin 1000 times, and record the sequence of outcomes.

What if the coin is biased \( P(X_j=1)=p, \ P(X_j=0)=1-p \) ?

Def. Let \( X \) be a discrete random variable with possible values \( t_1, t_2, t_3, \ldots \).
Expectation

Q: Is the expectation $E(X)$ the value that $X$ is equal to most often?

(a) Yes, always
(b) No, not generally

Example: Let $X$ be the number rolled on a fair die.

Example: Let $Y$ be $\text{Ber}(p)$. 
Expectation

Example You toss a biased coin repeatedly until the first heads. How long do you expect it to take?
Examples. Binomial

\[ S_n \sim \text{Bin}(n, p) \quad (S_n = X_1 + X_2 + \cdots + X_n \text{ for } X_j \text{ independent } \text{Ber}(p)) \]

\[ E(S_n) = \]
Examples. Poisson

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

\[ E(X) = \]

Example. A factory has, on average, 3 accidents per month. Estimate the probability that there will be exactly 2 accidents this month.
Examples

Toss a fair coin until tails comes up. If this is on the first toss, you win 2 dollars and stop. If heads comes up, the pot doubles and you continue. That is, if the first tails is on the k-th toss, you win $2^k$ dollars. What is your expected winnings?
Expectation of continuous random variables

$X$ discrete, $X \in \{t_1, t_2, \ldots \}$

$E(X) = \sum_k t_k P(X = t_k)$

$X$ continuous $P(X = t) = 0$ for each $t \in \mathbb{R}$

with density $f_X(t)$

Example

Let $U \sim \text{Unif}([a, b])$, $f_U(t) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq t \leq b \\ 0 & \text{otherwise} \end{cases}$
Example

Q: Shoot an arrow at a circular target of radius 1. What is the expected distance of the arrow from the center?

a) 1

b) $\frac{2}{3}$

c) $\frac{1}{2}$

d) $\frac{1}{4}$

e) 0