Probability: Why do we care?

- Probability helps us by:
  - Allowing us to translate scientific questions into mathematical notation
  - Providing a framework for answering scientific questions
- Later, we will see how some common statistical methods in the scientific literature are actually probability concepts in disguise

What is Probability?

- Probability is a measure of uncertainty about the occurrence of events
- Two definitions of probability
  - Classical definition
  - Relative frequency definition

Classical Definition

- \( P(E) = \frac{m}{N} \)
- If an event can occur in \( N \) equally likely and mutually exclusive ways, and if \( m \) of these ways possess the characteristic \( E \), then the probability of \( E \) is \( \frac{m}{N} \)
Example: Coin toss

- Flip one coin
- Tails and heads equally likely
- $N = 2$ possible events
- Let $H =$ Heads and $T =$ Tails

We are interested in the probability of tails: $P(Tails) = P(T) = \frac{1}{2}$

Relative Frequency Definition

- $P(E) = \frac{m}{n}$
- If an experiment is repeated $n$ times, and characteristic $E$ occurs $m$ of those times, then the relative frequency of $E$ is $\frac{m}{n}$, and it is approximately equal to the probability of $E$

Example: Multiple coin tosses I

Flip 100 coins

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>T = Tails</td>
<td>53</td>
</tr>
<tr>
<td>H = Heads</td>
<td>47</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
</tr>
</tbody>
</table>

$P(Tails) = P(T) \approx \frac{53}{100} = 0.53 \approx 0.50$

Example: Multiple coin tosses II

What happens if we flip 10,000 coins?

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>T = Tails</td>
<td>5063</td>
</tr>
<tr>
<td>H = Heads</td>
<td>4937</td>
</tr>
<tr>
<td>Total</td>
<td>10000</td>
</tr>
</tbody>
</table>

$P(Tails) = P(T) \approx \frac{5063}{10000} = 0.51 \approx 0.50$
The probability of $T$ is the limit of the relative frequency of $T$, as the sample size $n$ goes to infinity.

“The long run relative frequency”

Statistical independence

Two events are **statistically independent** if the joint probability of both events occurring is the product of the probabilities of each event occurring:

$$P(A \text{ and } B) = P(A) \times P(B)$$

Example

- Let $A =$ first born child is female
- Let $B =$ second child is female
- $P(A \text{ and } B) =$ probability that first and second children are both female:
- Assuming independence:

$$P(A \text{ and } B) = P(A) \times P(B) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$
“In a study where we are selecting patients at random from a population of interest, we assume that the outcomes we observe are independent...”

In what situations would this assumption be violated?

Mutually exclusive

- Two events are mutually exclusive if the joint probability of both events occurring is 0:
  \[ P(A \text{ and } B) = 0 \]

- Ex: A = first child is female, B = first child is male

ASIDE: Independence and Mutual exclusivity aren’t the same

- Two events are independent if the joint probability of both events occurring is the product of the probabilities of each event occurring:
  \[ P(A \text{ and } B) = P(A) \times P(B) \]

- Two events are mutually exclusive if the joint probability of both events occurring is 0:
  \[ P(A \text{ and } B) = 0 \]

- So the events A and B are both mutually exclusive AND independent only when \( P(A) = 0 \) or \( P(B) = 0 \).

Probability rules

1. The probability of any event is non-negative, and no greater than 1:
   \[ 0 \leq P(E) \leq 1 \]
2. Given \( n \) mutually exclusive events, \( E_1, E_2, \ldots, E_n \) covering the sample space, the sum of the probabilities of events is 1:
   \[ \sum_{i=1}^{n} P(E_i) = P(E_1) + P(E_2) + \cdots + P(E_n) = 1 \]
3. If \( E_i \) and \( E_j \) are mutually exclusive events, then the probability that either \( E_i \) or \( E_j \) occur is:
   \[ P(E_i \cup E_j) = P(E_i) + P(E_j) \]
Set notation

- A set is a group of disjoint objects
- An element of a set is an object in the set
- The union of two sets, A and B, is a larger set that contains all elements in either A, B or both
  Notation: \( A \cup B \)
- The intersection of two sets, A and B, is the set containing all elements found in both A and B
  Notation: \( A \cap B \)

The addition rule

If two events, A and B, are not mutually exclusive, then the probability that event A or event B occurs is:
\[
P(A \cup B) = P(A) + P(B) - P(A \cap B)
\]
where \( P(A \cap B) \) is the probability that both events occur

The multiplication rule

- In general:
  \[
P(A \cap B) = P(B) \times P(A|B)
  \]
- When events A and B are independent, \( P(A|B) = P(A) \) and:
  \[
P(A \cap B) = P(A) \times P(B)
  \]
Bayes rule

- Useful for computing $P(B|A)$ if $P(A|B)$ and $P(A|B^c)$ are known
- Ex: Screening
  - We know $P(\text{test positive} \mid \text{true positive})$
  - We want $P(\text{true positive} \mid \text{test positive})$
- Ex: Bayesian statistics uses assumptions about $P(\text{data} \mid \text{state of the world})$ to derive statements about $P(\text{state of the world} \mid \text{data})$

- The rule:

  $$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A|B) \cdot P(B) + P(A|B^c) \cdot P(B^c)}$$

  where $B^c$ denotes “the complement of B” or “not B”

---

Example: Sex and Age I

<table>
<thead>
<tr>
<th>Age</th>
<th>Young ($B_1$)</th>
<th>Older ($B_2$)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male ($A_1$)</td>
<td>30</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>Female ($A_2$)</td>
<td>40</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>Total</td>
<td>70</td>
<td>30</td>
<td>100</td>
</tr>
</tbody>
</table>

$P(A_1) = P(\text{male}) = \frac{50}{100} = 0.5$

$P(A_2) = P(\text{female}) = \frac{50}{100} = 0.5$

$P(B_1) = P(\text{young}) = \frac{70}{100} = 0.7$

$P(B_2) = P(\text{older}) = \frac{30}{100} = 0.3$

---

Example: Sex and Age II

<table>
<thead>
<tr>
<th>Age</th>
<th>Young ($B_1$)</th>
<th>Older ($B_2$)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male ($A_1$)</td>
<td>30</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>Female ($A_2$)</td>
<td>40</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>Total</td>
<td>70</td>
<td>30</td>
<td>100</td>
</tr>
</tbody>
</table>

$A_1 = \{\text{all males}\}$, $A_2 = \{\text{all females}\}$

$B_1 = \{\text{all young}\}$, $B_2 = \{\text{all older}\}$

$A_1 \cup A_2 = \{\text{all people}\} = B_1 \cup B_2$

$A_1 \cap A_2 = \{\text{no people}\} = \emptyset = B_1 \cap B_2$

$A_1 \cup B_1 = \{\text{male or young}\}$

$A_1 \cup B_2 = \{\text{male or old}\}$

$A_2 \cap B_2 = \{\text{female and old}\}$
Example: Sex and Age IV

<table>
<thead>
<tr>
<th>Age</th>
<th>Young ($B_1$)</th>
<th>Older ($B_2$)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male ($A_1$)</td>
<td>30</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>Female ($A_2$)</td>
<td>40</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>Total</td>
<td>70</td>
<td>30</td>
<td>100</td>
</tr>
</tbody>
</table>

\[
P(A_2 \cap B_2) = P(\text{older and female}) = \frac{10}{100} = 0.1
\]

\[
P(A_1 \cup B_1) = P(\text{young or male})
\]

\[
= P(A_1) + P(B_1) - P(A_1 \cap B_1)
\]

\[
= \frac{50}{100} + \frac{70}{100} - \frac{30}{100}
\]

\[
= \frac{90}{100} = 0.9
\]

Example: Sex and Age V

<table>
<thead>
<tr>
<th>Age</th>
<th>Young ($B_1$)</th>
<th>Older ($B_2$)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male ($A_1$)</td>
<td>30</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>Female ($A_2$)</td>
<td>40</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>Total</td>
<td>70</td>
<td>30</td>
<td>100</td>
</tr>
</tbody>
</table>

\[
P(B_2|A_2) = P(\text{older|female})
= \frac{P(B_2 \cap A_2)}{P(A_2)} = \frac{10/100}{50/100} = \frac{10}{50} = 0.2
\]

\[
P(B_2|A_1) = P(\text{older|male})
= \frac{P(B_2 \cap A_1)}{P(A_1)} = \frac{20/100}{50/100} = \frac{20}{50} = 0.4
\]

\[
P(B_2) = P(\text{older}) = \frac{30}{100} = 0.3
\]

Example: Sex and Age VI

\[
P(B_2|A_2) \neq P(B_2|A_1) \neq P(B_2)
\]

→ In this group, sex and age are not independent

Example: Sex and Age VII

<table>
<thead>
<tr>
<th>Age</th>
<th>Young ($B_1$)</th>
<th>Older ($B_2$)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male ($A_1$)</td>
<td>30</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>Female ($A_2$)</td>
<td>40</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>Total</td>
<td>70</td>
<td>30</td>
<td>100</td>
</tr>
</tbody>
</table>

Try these on your own...

\[
P(A_1 \cup A_2) =
\]

\[
P(B_1 \cup B_2) =
\]

\[
P(A_2|B_2) =
\]
Example: Blood Groups I

<table>
<thead>
<tr>
<th>Blood group</th>
<th>Sex</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>O</td>
<td>113</td>
<td>170</td>
</tr>
<tr>
<td>A</td>
<td>103</td>
<td>155</td>
</tr>
<tr>
<td>B</td>
<td>25</td>
<td>37</td>
</tr>
<tr>
<td>AB</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>251</td>
<td>377</td>
</tr>
</tbody>
</table>

Example: Blood Groups II

\[
P(\text{male}) = 1 - P(\text{female}) = \frac{251}{628} \approx 0.4
\]

\[
P(O) = \frac{283}{628} \approx 0.45
\]

\[
P(A) = \frac{258}{628} \approx 0.41
\]

\[
P(B) = \frac{62}{628} \approx 0.10
\]

\[
P(AB) = \frac{25}{628} \approx 0.04
\]

Example: Blood Groups III

Question: Are sex and blood group independent?

\[
P(O|\text{male}) = \frac{113}{251} \approx 0.45
\]

\[
P(O|\text{female}) = \frac{170}{377} \approx 0.45
\]

same as \(P(O) = \frac{283}{628} \approx 0.45\)

Can show same equalities for all blood types

→ Yes, sex and blood group appear to be independent of each other in this sample

Example: Disease in the population

- For patients with Disease X, suppose we knew the age proportions per sex, as well as the sex distribution.
- Question: Could we compute the sex proportions in each age group (young / older)?
- Answer: Use Bayes Rule

- \(A_1 = \{\text{males}\}, A_2 = \{\text{females}\}\)
- \(B_1 = \{\text{young}\}, B_2 = \{\text{older}\}\)

\[
P(A_1) = P(A_2) = 0.5
\]

\[
P(B_1 | A_2) = 0.2
\]

\[
P(B_2 | A_1) = 0.4
\]

\[
P(A_2 | B_2) = \frac{P(B_2 | A_2) \cdot P(A_2)}{P(B_2 | A_2) \cdot P(A_2) + P(B_2 | A_1) \cdot P(A_1)}
\]
Probability Distributions

- Often, we assume a true underlying distribution
  - Ex: $P(\text{tails}) = \frac{1}{2}$, $P(\text{heads}) = \frac{1}{2}$
- This distribution is characterized by a mathematical formula and a set of possible outcomes
- Two types of distributions:
  - Discrete
  - Continuous

Commonly Used Distributions

Discrete
- Binomial – two possible outcomes
  - Underlies much of statistical applications to epidemiology
  - Basic model for logistic regression
- Poisson – uses counts of events at rates
  - Basis for log-linear models

Continuous
- Normal – bell shaped curve
  - Many characteristics are normally distributed or approximately normally distributed
  - Basic model for linear regression
- Exponential – useful in describing growth

Counting techniques

- **Factorials**: count the number of ways to arrange things
- **Permutations**: count the number of possible ordered arrangements of subsets of a given size
- **Combinations**: count the number of possible unordered arrangements of subsets of a given size

Factorials

- Notation: $n!$ (“n factorial”)
- Number of possible arrangements of $n$ objects
- $n! = n(n-1)(n-2)(n-3) \cdots (3)(2)(1)$
- Standard convention: $0! = 1$
The number of ways you can take $r$ objects from a total of $n$ objects when order matters

$$nP_r = \frac{n!}{(n-r)!} = \frac{n(n-1)\cdots(n-r+1)(n-r)\cdots1}{(n-r)(n-r-1)\cdots1} = n(n-1)(n-2)\cdots(n-r+1)$$

The number of ways you can take $r$ objects from a total of $n$ objects when order doesn’t matter

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

For example:

$$\binom{4}{2} = \frac{4!}{2!(4-2)!} = \frac{4 \cdot 3 \cdot 2 \cdot 1}{2! \cdot 2!} = \frac{12}{2} = 6$$

Note: the number of combinations is less than or equal to the number of permutations.

You’ve seen it before:
- 2 x 2 tables and applications
- Proportions: CIs and tests
- Sensitivity and Specificity
- Odds ratio and relative risk
- Logistic regression

Bernoulli trial model
- The study of experiment consists of $n$ smaller experiments (trials) each of which has only two possible outcomes
  - Dead or alive
  - Success of failure
  - Diseased, not diseased
- The outcomes of the trials are independent
- The probabilities of the outcomes of the trial remain the same from trial to trial
Binomial Distribution Function

The probability of obtaining x “successes” in n Bernoulli trials is:

\[ P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} \]

where:
- \( p \) = probability of a “success
- \( q = 1-p \) = probability of “failure”
- \( X \) is a random variable
- \( x \) is a particular number

---

Example: Binomial (n=2) I

What is the probability, in a random sample of size 2, of observing 0, 1, or 2 heads?

<table>
<thead>
<tr>
<th># heads</th>
<th>Possible outcome</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>HH</td>
<td>( p \cdot p = p^2 )</td>
</tr>
<tr>
<td>1</td>
<td>HT</td>
<td>( p \cdot q )</td>
</tr>
<tr>
<td>1</td>
<td>TH</td>
<td>( q \cdot p )</td>
</tr>
<tr>
<td>0</td>
<td>TT</td>
<td>( q \cdot q = q^2 )</td>
</tr>
</tbody>
</table>

---

Example: Binomial (n=2) II

Recall \( x \) = number of observed heads

\[ P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} \]

\[
egin{align*}
P(X = 0) &= \binom{2}{0} (0.5)^0 (0.5)^2 = 0 \\
&= 0.25 = q^2
\end{align*}
\]

---

Example: Binomial (n=2) III

\[ P(X = 1) = \binom{2}{1} (0.5)^1 (0.5)^2 - 1 \]

\[
egin{align*}
&= \frac{2!}{1!(2 - 1)!} (0.5)(0.5) \\
&= 2(0.5)(0.5) = 0.5 = 2 \cdot p \cdot q
\end{align*}
\]

\[ P(X = 2) = \binom{2}{2} (0.5)^2 (0.5)^0 \]

\[
egin{align*}
&= \frac{2!}{2!(2 - 2)!} (0.5)^2 (0.5)^0 \\
&= 0.25 = p^2
\end{align*}
\]
Example: Binomial (n=3) I

<table>
<thead>
<tr>
<th># successes</th>
<th>Samples</th>
<th>P(X=x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>{+++}</td>
<td>(_{3}^{3}p^3q^0 = p^3)</td>
</tr>
<tr>
<td>2</td>
<td>{++-, +++, -++}</td>
<td>(_{3}^{2}p^2q = 3p^2q)</td>
</tr>
<tr>
<td>1 Nov</td>
<td>{-+-, -+-, --+}</td>
<td>(_{3}^{1}pq^2 = 3pq^2)</td>
</tr>
<tr>
<td>0</td>
<td>{---}</td>
<td>(_{3}^{0}p^0q^3 = q^3)</td>
</tr>
</tbody>
</table>

Example: Binomial (n=3) II

Since \(X\) takes discrete values only:

\[
P(X \leq 1) = P(X = 0) + P(X = 1) \\
P(X < 1) = P(X = 0) \\
P(X > 2) = P(X = 3) \\
P(1 \leq X \leq 2) = P(X = 1) + P(X = 2) \\
P(X \geq 1) = P(X = 1) + P(X = 2) + P(X = 3) = 1 - P(X = 0)
\]

Example: Binomial (n=3) III

The probability that a person suffering from a head cold will obtain relief with a particular drug is 0.9. Three randomly selected sufferers from the cold are given the drug.

- \(p = 0.9\)
- \(q = 1-p = 0.1\)
- \(n = 3\)

Example: Binomial (n=3) IV

\(2^3 = 8\) possible outcomes

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>S S S</td>
<td>ppp</td>
</tr>
<tr>
<td>S S F</td>
<td>ppq</td>
</tr>
<tr>
<td>S F S</td>
<td>pqq</td>
</tr>
<tr>
<td>F S S</td>
<td>qpp</td>
</tr>
<tr>
<td>F F S</td>
<td>qpp</td>
</tr>
<tr>
<td>F S F</td>
<td>qpp</td>
</tr>
<tr>
<td>S F F</td>
<td>pqq</td>
</tr>
<tr>
<td>F F F</td>
<td>qqq</td>
</tr>
</tbody>
</table>
Example: Binomial (n=3) V

Probability exactly zero (none) obtain relief:

\[ P(X = 0) = \binom{3}{0} p^0 q^3 = q^3 \]
\[ = (0.1)^3 = 0.001 \]

Probability exactly one obtains relief:

\[ P(X = 1) = \binom{3}{1} p^1 q^2 \]
\[ = \frac{3!}{1!2!} p q^2 = \frac{3 \cdot 2!}{1 \cdot 2!} p q^2 = 3pq^2 \]
\[ = 3(0.9)(0.1)^2 = 0.027 \]

Example: Bernoulli Distribution

Let \( X = 1 \) with probability \( p \), and 0 otherwise

Calculation of the mean:

\[ E(X) = \mu = \sum_{i=0,1} x_i P(X = x_i) = 1 \cdot p + 0 \cdot (1 - p) = p \]

Calculation of the variance:

\[ E(X^2) = (1^2) \cdot p + (0^2) \cdot (1 - p) = p \]
\[ Var(X) = E(X^2) - \mu^2 = p - p^2 = p \cdot (1 - p) \]

Mean and Variance

Mean of a random variable (r.v.) \( X \)

- Expected value, expectation
- \( \mu = E(X) \)
- \( \sum_i x_i P(X = x_i) \) for discrete r.v.
- \( \int_{-\infty}^{+\infty} x \cdot f(x) \, dx \) for continuous r.v.

Variance of a random variable, \( X \)

- \( \sigma^2 = Var(X) = E(X - \mu)^2 = E(X^2) - \mu^2 \)
- The standard deviation \( \sigma = \sqrt{\sigma^2} = \sqrt{Var(X)} \)

Properties of Expectation

1. \( E(c) = c \) where \( c \) is a constant
2. \( E(c \cdot X) = c \cdot E(X) \)
3. \( E(X_1 + X_2) = E(X_1) + E(X_2) \)
**Properties of Variance**

1. \( \text{Var}(c) = 0 \) where \( c \) is a constant
2. \( \text{Var}(c \cdot X) = c^2 \cdot \text{Var}(X) \)
3. \( \text{Var}(X_1 + X_2) = \text{Var}(X_1) + \text{Var}(X_2) \) if \( X_1 \) and \( X_2 \) are independent

**Binomial Mean and Variance**

- \( S \) is Binomial \((n, p)\), so...
- \( S = \sum_{i=1}^{n} X_i \) where \( X_i \) are independent Bernoulli\((p)\) random variables
- \( E(S) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} p = np \)
- \( \text{Var}(S) = \sum_{i=1}^{n} \text{Var}(X_i) = \sum_{i=1}^{n} p(1 - p) = np(1 - p) \)

**Poisson Distribution**

- Describes occurrences or objects which are distributed randomly in space or time
- Often used to describe distribution of the number of occurrences of a rare event
- Underlying assumptions similar to those for binomial distribution
- Useful when there are counts with no denominator

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Parameters needed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binomial</td>
<td>( n, p )</td>
</tr>
<tr>
<td>Poisson</td>
<td>( \lambda = np )</td>
</tr>
<tr>
<td></td>
<td>= the expected number of events per unit time</td>
</tr>
</tbody>
</table>

**Poisson Distribution Examples**

- Number of Prussian officers killed by horse kicks between 1875 and 1894
- Spatial distribution of stars, weeds, bacteria, flying-bomb strikes
- Emergency room or hospital admissions
- Typographical errors
- Deaths due to a rare disease
Poisson Assumptions

- The occurrences of a random event in an interval of time are independent.
- In theory, an infinite number of occurrences of the event are possible (though perhaps rare) within the interval.
- In any extremely small portion of the interval, the probability of more than one occurrence of the event is approximately zero.

Poisson Probability

- The probability of $x$ occurrence of an event in an interval is:
  \[ P(X = x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}, \quad x = 0, 1, 2, \ldots \]
  where $\lambda$ is the expected number of occurrences in the interval.
- $e$ is a constant ($\approx 2.718$).
- For the Poisson distribution, mean = variance = $\lambda$.

Example: Traffic accidents I

- Suppose the goal has been set of bringing the expected number of traffic accidents per day in Baltimore down to 3. There are 5 fatal accidents today. Has the goal been attained?
- The number of accidents follows a Poisson distribution because:
  - The population that drives in Baltimore is large.
  - The number of accidents is relatively small.
  - People have similar risks of having an accident.
  - The number of people driving each day is fairly stable.
  - The probability of two accidents occurring at exactly the same time is approximately zero.

Example: Traffic accidents II

- We are aiming for a rate of $\lambda = 3$ fatal accidents per day, or lower.
- The observed number is 5.
- $P(X = 5; \lambda = 3) = \frac{e^{-3} \cdot 3^5}{5!} = 0.101$.
- Has the goal been attained?
Example: Suicide in the City

If the rate for a given rare condition is expressed as $\mu$ per time period, the expected number of events is $\mu t$ where $t$ is the time period.

Some questions we can answer using the properties of the Poisson distribution are:

- Suppose the weekly rate of suicide in a large city is 2. What is the probability of one suicide in a given week?
  
  $P(X = 1; \lambda = 2)$

- What is the probability of 2 suicides in 2 weeks?

- Since the weekly rate of suicide was 2/week, we expect $2 \times 2$ or 4 suicides per 2 week period.

- You can use the poisson distribution to calculate $P(X = 2; \lambda = 4)$

Example: Cancer in a large population

Yearly cases of esophageal cancer in a large city; 30 cases observed in 1990

$P(X = 30) = \frac{e^{-\lambda} \lambda^{30}}{30!}$

where $\lambda =$ yearly average number of cases of esophageal cancer

Example: Down’s syndrome I

The incidence of Down’s syndrome as a function of mother’s age

The Poisson distribution can be used to approximate a Binomial($n,p$) distribution when:

- $n$ is large and $p$ is very small, or
- $np = \lambda$ is fixed, and $n$ becomes infinitely large
Example: Down’s syndrome II

- Suppose the incidence of Down’s syndrome in 40-year-old mothers is 1/100
- Out of 25 babies born to 40-year-old women, what is the frequency of babies with Down’s syndrome?
- We can approach this problem using a Binomial(25, 1/100) model, or using a Poisson($\lambda = 0.25$) model

Example: Down’s syndrome III

<table>
<thead>
<tr>
<th>Babies with Down’s Syndrome</th>
<th>$P(X=x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Poisson</td>
</tr>
<tr>
<td>0</td>
<td>0.779</td>
</tr>
<tr>
<td>1</td>
<td>0.195</td>
</tr>
<tr>
<td>2</td>
<td>0.024</td>
</tr>
<tr>
<td>&gt;2</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Note: the approximation becomes even better for larger values of $n$.

Lecture 2 Summary

We’ve covered a lot of ground this lecture. Here’s a quick summary of what we just discussed:

- Probability
  - Commonly used definitions and properties including: independence, mutually exclusive, addition rule, conditional probability, the multiplication rule and Bayes rule
- Discrete distributions
  - Binomial and Poisson

Tomorrow we’ll discuss the Normal distribution, the Central Limit Theorem, the t-distribution and confidence intervals.