Outline

1. Poisson distribution
2. Tests of hypothesis for a single Poisson mean
3. Comparing multiple Poisson means
4. Likelihood equivalence with exponential model
## Pump failure data

<table>
<thead>
<tr>
<th>Pump</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Failures</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>14</td>
<td>3</td>
<td>19</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>22</td>
</tr>
<tr>
<td>Time</td>
<td>94.32</td>
<td>15.72</td>
<td>62.88</td>
<td>125.76</td>
<td>5.24</td>
<td>31.44</td>
<td>1.05</td>
<td>1.05</td>
<td>2.10</td>
<td>10.48</td>
</tr>
</tbody>
</table>
The Poisson distribution

- Used to model counts
- The Poisson mass function is

\[ P(X = x; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \]

for \( x = 0, 1, \ldots \)

- The mean of this distribution is \( \lambda \)
- The variance of this distribution is \( \lambda \)
- Notice that \( x \) ranges from 0 to \( \infty \)
Some uses for the Poisson distribution

- Modeling event/time data
- Modeling radioactive decay
- Modeling survival data
- Modeling unbounded count data
- Modeling contingency tables
- Approximating binomials when \( n \) is large and \( p \) is small
Definition

- $\lambda$ is the mean number of events per unit time
- Let $h$ be very small
- Suppose we assume that
  * Prob. of an event in an interval of length $h$ is $\lambda h$ while the prob. of more than one event is negligible
  * Whether or not an event occurs in one small interval does not impact whether or not an event occurs in another small interval
then, the number of events per unit time is Poisson with mean $\lambda$
Poisson approximation to the binomial

- When $n$ is large and $p$ is small the Poisson distribution is an accurate approximation to the binomial distribution.

- Notation
  
  * $\lambda = np$
  
  * $X \sim \text{Binomial}(n, p)$, $\lambda = np$ and
  
  * $n$ gets large
  
  * $p$ gets small
  
  * $\lambda$ stays constant
**Proof** Rice page 41

\[ P(X = k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \]

\[ = \frac{n!}{k!(n-k)!} \left( \frac{\lambda}{n} \right)^k \left( 1 - \frac{\lambda}{n} \right)^{n-k} \]

\[ = \frac{n!}{(n-k)!n^k} \times \left( 1 - \frac{\lambda}{n} \right)^{-k} \times \frac{\lambda^k}{k!} \left( 1 - \frac{\lambda}{n} \right)^n \]

\[ \rightarrow 1 \times 1 \times \frac{\lambda^k}{k!} e^{-\lambda} \]

\[ = \frac{\lambda^k}{k!} e^{-\lambda} \]
Notes

• That \((1 - \lambda/n)^n\) converges to \(e^{-\lambda}\) for large \(n\) is a very old mathematical fact

• We can show that \(\frac{n!}{(n-k)!n^k}\) goes to one easily because

\[
\frac{n!}{(n-k)!n^k} = 1 \times \left(1 - \frac{1}{n}\right) \times \left(1 - \frac{2}{n}\right) \times \ldots \times \left(1 - \frac{k-1}{n}\right)
\]

each term goes to 1
Examples
Some example uses of the Poisson distribution
  deaths per day in a city
  homicides witnessed in a year
  teen pregnancies per month
  Medicare claims per day
  cases of a disease per year
  cars passing an intersection in a day
  telephone calls received by a switchboard in an hour
Some results

• If $X \sim \text{Poisson}(t\lambda)$ then

$$\frac{X - t\lambda}{\sqrt{t\lambda}} = \frac{X - \text{Mean}}{\text{SD}}$$

converges to a standard normal as $t\lambda \to \infty$

• Hence

$$\frac{(X - t\lambda)^2}{t\lambda} = \frac{(O - E)^2}{E}$$

converges to a Chi-squared with 1 degree of freedom for large $t\lambda$

• If $X \sim \text{Poisson}(t\lambda)$ then $X/t$ is the ML estimate of $\lambda$
Proof

- Likelihood is
  \[ \frac{(t\lambda)^x e^{-t\lambda}}{x!} \]

- So that the log-likelihood is
  \[ x \log(\lambda) - t\lambda + \text{constants in } \lambda \]

- The derivative of the log likelihood is
  \[ x/\lambda - t \]

- Setting equal to 0 we get that \( \hat{\lambda} = x/t \)
Pump failure data

• Failures for Pump 1: 5, monitoring time: 94.32 days
• Estimate of $\lambda$, the mean number of failures per day
  $= 5/94.32 = .053$
• Test the hypothesis that the mean number of failures per day is larger than the industry standard, .15 events per day: $H_0 : \lambda = .15$ versus $H_a : \lambda > .15$
• $TS = (5 - 94.32 \times .15)/\sqrt{94.2 \times .15} = -2.433$
• Hence P-value is very large (.99)
• HW: Obtain a confidence interval for $\lambda$
Pump failure data

• Exact P-value can be obtained by using the Poisson distribution directly

\[ P(X \geq 5) \text{ where } X \sim \text{Poisson}(0.15 \times 94.32) \]

\[
\text{ppois}(5, 0.15 \times 94.32, \text{lower.tail = FALSE}) = 0.995
\]

very little evidence to suggest that this pump is malfunctioning

• To obtain a P-value for a two-sided alternative, double the smaller of the two one sided P-values
<table>
<thead>
<tr>
<th>Pump</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Failures</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>Time</td>
<td>94.32</td>
<td>15.72</td>
<td>62.88</td>
<td>125.76</td>
<td>5.24</td>
</tr>
<tr>
<td>$\hat{\lambda}$</td>
<td>0.053</td>
<td>0.064</td>
<td>0.080</td>
<td>0.111</td>
<td>0.573</td>
</tr>
<tr>
<td>P-value</td>
<td>0.995</td>
<td>0.999</td>
<td>0.908</td>
<td>0.843</td>
<td>0.009</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pump</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Failures</td>
<td>19</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>22</td>
</tr>
<tr>
<td>Time</td>
<td>31.44</td>
<td>1.05</td>
<td>1.05</td>
<td>2.10</td>
<td>10.48</td>
</tr>
<tr>
<td>$\hat{\lambda}$</td>
<td>0.604</td>
<td>0.952</td>
<td>0.952</td>
<td>1.904</td>
<td>2.099</td>
</tr>
<tr>
<td>P-value</td>
<td>1e-7</td>
<td>0.011</td>
<td>0.011</td>
<td>1e-5</td>
<td>2e-19</td>
</tr>
</tbody>
</table>
Common failure rate

- If $X_i$ for $i = 1, \ldots, n$ are Poisson$(t_i \lambda)$ then
  $$\sum X_i \sim \text{Poisson} \left( \lambda \sum t_i \right)$$

  If you are willing to assume the common $\lambda$, then $\sum X_i$ contains all of the relevant information.

- Clearly a common $\lambda$ across pumps is not warranted. However, for illustration, assume that this is the case.

- Then the total number of failures was 75.

- The total monitoring time was 305.4.

- The estimate of the common $\lambda$ would be $75/305.4 = .246$. 
Person-time analysis

<table>
<thead>
<tr>
<th>OC</th>
<th># of cases</th>
<th># of person-years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current users</td>
<td>9</td>
<td>2,935</td>
</tr>
<tr>
<td>Never users</td>
<td>239</td>
<td>135,130</td>
</tr>
</tbody>
</table>

• One of the most common uses of the Poisson distribution in epi is to model rates

• Estimated incidence rate amongst current users is $\frac{9}{2,935} = 0.0038$ events per person year

• Estimated incidence rate amongst never users is $\frac{239}{135,130} = 0.0018$ events per person year
Poisson model

- Rates are often *modeled* as Poisson
- Notice that the total number of subjects is discarded, whether the 2,935 years was comprised of 1,000 or 500 people does not come into play
- Most useful for rare events (though we’ll discuss another motivation for the Poisson model later)

* $\lambda = .3$ events per year
* Followed 10 people for a total of 40 person-years
* Expected number of deaths is $0.3 \times 40 = 12$, larger than our sample

- $\lambda$ is assumed constant over time
Comparing two Poisson means

- Want to test $H_0 : \lambda_1 = \lambda_2 = \lambda$
- Observed counts $x_1, x_2$ Person-times $t_1, t_2$
- Estimate of $\lambda$ under the null hypothesis is $
  \hat{\lambda} = \frac{x_1 + x_2}{t_1 + t_2}$
- Estimated expected count in Group 1 under $H_0$
  \[ E_1 = \hat{\lambda}t_1 = (x_1 + x_2)\frac{t_1}{t_1 + t_2} \]
- Estimated expected count in Group 2 under $H_0$
  \[ E_2 = \hat{\lambda}t_2 = (x_1 + x_2)\frac{t_2}{t_1 + t_2} \]
Notes

• Test statistic

\[ TS = \sum \frac{(O - E)^2}{E} = \frac{(x_1 - E_1)^2}{E_1} + \frac{(x_2 - E_2)^2}{E_2} \]

follows a Chi-squared distribution with 1 df

• Equivalent computational form

\[ TS = \frac{(X_1 - E_1)^2}{V_1} \]

where

\[ V_1 = \frac{(x_1 + x_2)t_1t_2}{(t_1 + t_2)^2} \]
OC example

- \( x_1 = 9, \ t_1 = 2,935 \)
- \( x_2 = 239, \ t_2 = 135,130 \)
- \( E_1 = (9 + 239) \times \frac{2,935}{2,935 + 135,130} = 5.27 \)
- \( V_1 = (9 + 239) \times \frac{2,935 \times 135,130}{(2,935 + 135,130)^2} = 5.16 \)

\[
TS = \frac{(9 - 5.27)^2}{5.16} = 2.70
\]

P-value = .100
Estimating the relative rate

- Relative rate $\lambda_1/\lambda_2$
- Estimate $(x_1/t_1)/(x_2/t_2)$
- Standard error for the log relative rate estimate
  \[ \frac{1}{x_1} + \frac{1}{x_2} \]
- For the OC example the estimated log relative rate is .550
- The standard error is $\sqrt{\frac{1}{9} + \frac{1}{239}} = .340$
- 95% CI for the log relative rate is
  \[ .550 \pm 1.96 \times .340 = (−.115, 1.26) \]
Alternate motivation for Poisson model

- Another way to motivate the Poisson model is to show that it is *likelihood equivalent* to a plausible model.
- We show that the Poisson model is likelihood equivalent to a model that specifies failure times as being independent exponentials.
- Whenever the independent exponential model is reasonable, then so is the Poisson model, regardless of how large or small the rate or sample size is.
- The likelihood equivalence implies that likelihood methods apply.
Recall

- The likelihood is the density viewed as a function of the parameter.
- The likelihood summarizes the evidence in the data about the parameter.
- If $X \sim \text{Poisson}(t\lambda)$ then the likelihood for $\lambda$ is
  \[ L(\lambda) = \frac{(t\lambda)^x e^{-t\lambda}}{x!} \propto \lambda^x e^{-t\lambda} \]
- When you have independent observations, the likelihood is the product of the likelihood for each observation.
Likelihood equivalence with exponential model

• Suppose there is no censoring (every person followed until an event)

• We model each person’s time until failure as independent exponentials: \( Y_i \sim \text{Exponential}(\lambda) \)

• Each subject contributes \( \lambda e^{-y_i \lambda} \) to the likelihood

\[
\prod_{i=1}^{n} \lambda e^{-y_i \lambda} = \lambda^n \exp \left( -\lambda \sum y_i \right)
\]

• Here \( n \) is the number of events and \( \sum y_i \) is the total person-time

• Same likelihood as if we specify \( n \sim \text{Poisson}(\lambda \sum y_i) \)!
Likelihood equivalence with censoring

- Suppose now that the study ended before events were observed on some subjects.

- Likelihood contribution for each event is $\lambda e^{-y_i \lambda}$.

- Likelihood contribution for each censored observation is

\[ P(Y \geq y_i; \lambda) = \int_{y_i}^{\infty} \lambda e^{-u \lambda} du = e^{-y_i \lambda} \]

The contribution is this integral b/c we don’t really know the event time for that person, only that it was later than $y_i$. 
• Combining the censored and non-censored observations, we have
\[
\left\{ \prod_{\text{non-censored } i} \lambda e^{-y_i \lambda} \right\} \times \left\{ \prod_{\text{censored } i} e^{-y_i \lambda} \right\}
\]
x events (and hence \(n - x\) censored observations)
\[
= \lambda^x \exp \left( -\lambda \sum_{i=1}^{n} y_i \right)
\]
• \(x\) is the total number of events and \(\sum_{i=1}^{n} y_i\) is the total amount of person-time
• This is exactly the same likelihood as modeling \(X \sim \text{Poisson}(\lambda \sum_{i=1}^{n} y_i)\)