AMSI7: WEEK 4. CLASS 1

Poisson distribution (Continuation)
The Normal distribution
Monday April 20th, 2015
Example: The Poisson distribution

- Homicide deaths: In one year, there were 116 homicide deaths in Richmond, VA (Data from Richards, 1991).
- For a randomly selected day, find the probability that the number of homicide deaths is:
  - a) 0; b) 1; c) 2; d) 3; e) 4

**STEPS:**
- Find $\mu$
  \[
  \mu = \frac{\text{# homicide deaths in one year}}{\text{# of days in one year}} = \frac{116}{365} = 0.317 \ \text{deaths/day}
  \]
- Find $P(x)$ when $x=0,1,2,3,4$
Example (Cont.)

- Poisson distribution: \( P(x) = \frac{\mu^x e^{-\mu}}{x!} \)
- \( x=0: P(x) = \frac{\mu^0 e^{-\mu}}{0!} = e^{-0.3178} = 0.728 \)
- \( x=1: P(x) = \frac{\mu^1 e^{-\mu}}{1!} = 0.3178 \cdot e^{-0.3178} = 0.231 \)
- \( x=2: P(x) = \frac{\mu^2 e^{-\mu}}{2!} = \frac{0.3178^2 e^{-0.3178}}{2} = 0.0367 \)
- \( x=3: P(x) = \frac{\mu^3 e^{-\mu}}{3!} = \frac{0.3178^3 e^{-0.3178}}{6} = 0.0039 \)
- \( x=4: P(x) = \frac{\mu^4 e^{-\mu}}{4!} = \frac{0.3178^4 e^{-0.3178}}{24} = 0.0003 \)
Example (Cont.)

- Comparison with actual results:

<table>
<thead>
<tr>
<th># of Deaths</th>
<th>P(X) (Prob.)</th>
<th>Expected # of Days (*)</th>
<th>Actual # of Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.728</td>
<td>265.6</td>
<td>268</td>
</tr>
<tr>
<td>1</td>
<td>0.231</td>
<td>84.4</td>
<td>79</td>
</tr>
<tr>
<td>2</td>
<td>0.0367</td>
<td>13.4</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>0.0039</td>
<td>1.4</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0.0003</td>
<td>0.1</td>
<td>0</td>
</tr>
</tbody>
</table>

(*): This is the expected number of days with 0, 1, 2, 3, 4 homicides. This column is calculated by multiplying P(x) by 365 days.

Since the predicted frequencies with the Poisson model are close to the observed frequencies, we say that the Poisson model has a good fit to the observed data.
IMPORTANT CONCEPTS

- RANDOM VARIABLE
- PROBABILITY DISTRIBUTION
- DISCRETE RANDOM VARIABLE
- CONTINUOUS RANDOM VARIABLE
Normal distribution

• The normal distribution is a continuous probability distribution

Formula: \( y = f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2} \)

• Graph: Bell-shaped: Curve is symmetric around the mean \( \mu \)

• Distribution is determined by two parameters: the mean \( \mu \) and the standard deviation \( \sigma \)
Graph of a Normal Distribution

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$N(\mu, \sigma^2)$
Properties of the Continuous Probability Distributions

Remember for a **discrete** probability distribution:

1) \( \sum P(X) = 1 \)
2) \( 0 \leq P(X) \leq 1 \) for all individual \( X \) values
3) Graph is a probability histogram

For a **continuous** probability distribution:

1) The total area under the curve **must be** equal to 1
2) Every point on the curve has a vertical height of 0 or greater
3) Graph is called a **density curve**

**NOTE:** There is a correspondence between area and probability
Uniform distribution

• This is a continuous probability distribution

EXAMPLE: Assume Voltages vary between 6 Volts and 12 Volts, and all the voltages have the same possibilities of happening.

PROBLEMS:

1) Find the probability of Voltage being greater than 10 Volts.
2) Find the probability of Voltage between 6.5 and 8 Volts.
Uniform Distribution Example

TOTAL RECTANGLE AREA = (12-6) \times \frac{1}{6} = 1

f(Volts) IS A PROBABILITY DISTRIBUTION

ANSWERS TO PROBLEMS

1) Area of interest (Black) = Probability = (12-10) \times \frac{1}{6} = \frac{1}{3}

2) Area of interest (Purple) = Probability = (8-6.5) \times \frac{1}{6} = \frac{1.5}{6} = \frac{1}{4}
Standard Normal Distribution

- Is a Normal probability distribution with mean 0 and standard deviation of 1 (Total area under the curve is 1)
Comparing Data Sets

- Two different possible Normal distributions

- EXAMPLE: Heights of Adult Women and Men

<table>
<thead>
<tr>
<th>WOMEN</th>
<th>MEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu = 63.6$ in</td>
<td>$\mu = 69$ in</td>
</tr>
<tr>
<td>$\sigma = 2.5$ in</td>
<td>$\sigma = 2.8$ in</td>
</tr>
</tbody>
</table>
Comparing two normal distributions

- Same mean, different variances
- Different means, same variances
- Different means, different variances

EXAMPLE:
Given a z Score find a probability

- Tables are available when \( \mu=0 \) and \( \sigma=1 \) (Standard Normal Distribution)

EXAMPLES: Suppose that thermometer readings at the water freezing point are normally distributed with mean 0°C and a standard deviation equal to 1.0 °C

1) Find the probability that at the freezing point of water, readings are less than 1.58°C.

From Table A-2 area is 0.9429. There is a 94.29% chance that a randomly selected thermometer will have readings below 1.58 °C (or 94.29% of thermometers will have readings below 1.58 °C)
Example 1

- From Table A-2 area to the left of $z=1.58$ is 0.9426
Example 2

- Probability that readings are above -1.23 °C
From table A-2 area to the left of -1.23 is 0.1093. We want the area to the right. We calculate:
1 - 0.1093 = 0.8907
Example 3

- Probability of chosen thermometer reads between -2.00 °C and 1.50 °C

- From Table A-2, area to the left of $z=-2$ is 0.0228
- From Table A-2 area to the left of $z=1.5$ is 0.9332
- Required probability: $0.9332 - 0.0228 = 0.9104$