Lecture 4 (Ch-1)

We have been talking about data, and histograms of data. A histogram pertains to data.

A Huge and Tricky Concept

But there is something else that looks like a histogram, but it’s not: A DISTRIBUTION.

A dist. is a purely mathematical thing that has nothing to do with data. So, for now, forget data (and hists).

In statistics, distributions are used to represent the population, while histograms are used to describe the sample (data). Later, we are going to learn how to tell something about the pop from a sample by comparing the dist. with a hist. But, again, for now, think of them as completely unrelated.

Example: \( y \sim f(x) \sim e^{-\frac{1}{2}x^2} \)

Technically, this \( f(x) \) is not a distribution! You will see why, in a minute. But it’s good enough to make the important point that a dist. is a mathematical thing (i.e., a function), not a histogram, even though they look alike.
Here is the more precise definition of a distribution.

**Defn:** A distribution, $f(x)$, $p(x)$, must satisfy:

1) $f(x) \geq 0 \quad \parallel \quad p(x) \geq 0$

2) $\int_{-\infty}^{\infty} f(x) \, dx = 1 \quad \parallel \quad \sum_{x} p(x) = 1$

For $x = \text{categorical}$, $p(x)$ is called the prob. density function (pdf).
For $x = \text{categorical}$, $p(x)$ is called the prob. mass function (pmf).

Generally, $f(x)$ and $p(x)$ are called distributions.

**Example:** $f(x) = e^{-\frac{1}{2}x^2}$, $-\infty < x < \infty$, is not a dist, because $\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} \, dx$ remind you of how to do such integrals: $= \sqrt{2\pi} \neq 1$

So, $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ is a dist. (also $f(x) \geq 0 \checkmark$)

**Example:** $f(x) = k \chi^8 (-x)$, $0 < x < 1$ dist?

$\int_{-\infty}^{\infty} f(x) \, dx = \int_{0}^{1} k \chi^8 (-x) \, dx = k \cdot \frac{1}{90} \neq 1$ unless $k = 90$.

So, $f(x) = 90 \chi^8 (-x)$ is a dist. (also $f(x) \geq 0 \checkmark$)

The variety of shapes that dists have is similar to that of hists (see previous lecture).
Recall that areas under histograms \( \sim \) proportion of times \( \sim \) probability based on data i.e. for sample.

Similarly, areas under distributions \( \sim \) proportion of times \( \sim \) probability based on distribution i.e. for population.

<table>
<thead>
<tr>
<th>Proportion/Probability of ( a &lt; x &lt; b )</th>
<th>Proportion/Probability of ( x = { \ldots, } )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \int_{a}^{b} f(x) , dx )</td>
<td>( \sum_{x={ \ldots, }} p(x) )</td>
</tr>
<tr>
<td>( f(x) )</td>
<td>( p(\text{Mac}) + p(\text{Dell}) + \ldots )</td>
</tr>
<tr>
<td>( a )</td>
<td>( x )</td>
</tr>
<tr>
<td>( b )</td>
<td>( \text{Mac, Dell,} \ldots )</td>
</tr>
</tbody>
</table>

E.g. for \( f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \)

\[ \text{prop}(a < x < b) = \int_{a}^{b} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \, dx = \text{whatever} \]

\[ \text{or prob.} \]

E.g. for \( f(x) = 90 \, x^8 \, (1-x) \)

\[ \text{prop}(a < x < b) = \int_{a}^{b} 90 \, x^8 \, (1-x) \, dx = \text{whatever} \]

\[ \text{or prob.} \]

\[ \text{Note:} \quad \text{prop}(x = \text{anything}) = 0 \quad \text{if} \quad x = \text{continuous.} \]

\[ \text{even at a line.} \]
More examples of dists:

- \( x = \text{cont.} \): Easier
  \[ f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} x^2} \quad x < \infty \quad \int_{-\infty}^{\infty} f(x)dx = 1 \]

- \( x = \text{categorical}. \) Harder

\[ \begin{array}{c|c|c|c}
 x & \text{Mac} & \text{Dell} & \text{HP} \\
 p(x) & 0.2 & 0.1 & 0.7 \\
\end{array} \]

\[ \sum p(x) = 1 \]

- Example:
  - \( x = \) “state of a fair coin”
    \[ p(x) \]
    \[ \begin{array}{c|c|c}
 x & H & T \\
 p(x) & 0.5 & 0.5 \\
\end{array} \]

- Example:
  - \( x = \) “number of heads out of \( n \) tosses of a fair coin.”

\[ p(x) = \frac{n!}{x!(n-x)!} \left( \frac{1}{2} \right)^x \left( \frac{1}{2} \right)^{n-x} \]

- Binomial distn.

We will derive this later, we will replace \( \frac{1}{2} \) with other values.

Note: There is no data anywhere here. These are not histograms or formulas.

Don’t forget; all these \( p(x) \)’s are used to describe the population of \( x \).
Q1: I mentioned the Bernoulli dist. above. Its pmf is defined by \( p(x) = \pi^x (1-\pi)^{1-x} \) where \( \pi = 0,1 \). What is the probability of getting \( x = 1 \)?

\( 0 \leq \pi \leq 1 \)

a) 0  b) 1  c) \( \pi \)  d) Insufficient info given.

\[ \text{prob}(x=1) = p(x=1) = \pi^1 (1-\pi)^{0} = \pi \]

When poll is active, respond at PollEv.com/marzban

Text MARZBAN to 22333 once to join

---

hw_lect4-1

Consider the density function \( f(x) = \frac{1}{3} (1-x^3 + x^2 + x + 2) \), \( 0 < x < 2 \)

a) First, determine \( a \) to make sure \( f(x) \) is a density function.

b) Compute the prob. That \( x \) will be between 0 and 1.

c) Use R to plot \( f(x) \). Include code and figure.

This problem is basically an exercise in calculus.

---

hw_lect4-2

The Bernoulli dist, discussed in the lecture does have a formula: \( p(x) = \pi^x (1-\pi)^{1-x} \), where \( 0 < \pi < 1 \) is some param. and \( x = 0,1 \).

Show that it's a mass function.