Probability, Statistics, and beyond ...

Igor Cialenco, IIT
igor@math.iit.edu, Office E1 234C
IIT, November 8, 2007
Stochastics is fun to do !!!

Stochastics is important for many real-life applications
What is Probability and Statistics

- **Probability** - random, chaotic, stochastic, uncertain. To find what are the chances of something to happen

- **Statistics** - \((\text{Probability})^{-1}\)

- **Stochastics** = Probability + Statistics + Applications
Flip a **fair** coin

All possible results or **outcomes** are **Head**, **Tail**.

**Sample Space** $\Omega = \{H, T\}$

Every outcome has equal chances to happen (the coin is fair!).

So, the chances are 50/50 for $H, T$

50% chances for $H$, and 50% for $T$

In formulas $P(H) = P(T) = \frac{1}{2} = .5$
A fair die \( \Rightarrow \) \( \Omega = \{1, 2, 3, 4, 5, 6\} \)

\[ P(k) = \frac{1}{6}, \text{ for } k = 1, \ldots, 6 \]

What are the chances (probabilities) of getting an even number?

\[ P(2, 4, 6) = \frac{3}{6} = \frac{1}{2} \]

\[ P(\text{not to have } 2, 5) = P(\text{to get } 1, 3, 4, 6) = \frac{4}{6} = \frac{2}{3} \]

In general:

if the outcomes are equally likely then for any event \( A \subset \Omega \)

\[ P(A) = \frac{\# \text{ elements in } A}{\# \text{ of total possible outcomes}} \]
Two dice

- $\Omega = \{(i, j) : i, j = 1, 2, 3, 4, 5, 6\}$ an array $6 \times 6$

- What is the probability that the sum is 7?

- What is the probability of getting 6?

- How about the probability that sum is 7 given that one die was 6?

- **Conditional probability.** $P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$.

- **HW:** A family has two children. What is the probability that both are boys, given that at least one is a boy? (Hint: the chances that one child is a Boy or a Girl is .5)
If the event $B$ does not affect the probability of event $A$, then independent

- $P(A) = P(A|B)$ or $P(A \text{ and } B) = P(A)P(B)$

- Ex: 3 coins. $A = \text{(the first coin is } H)$. $B = \text{(the third coin is } T)$. Independent.

- **HW**: Find some sets $A, B$ from the above example such that $A, B$ are dependent

- Ex: A box with color balls, say 5 Green and 3 Blue.
What can we do if the coin is not a fair one? How to detect that the coin is not fair? How unfair is it?

Specify the probabilities!

\[ P(H) = p, \text{ then } P(T) = 1 - p \]

In general, \( P \) is a probability if:

i) for every \( A \subset \Omega \), \( P(A) \geq 1 \)

ii) \( p(\Omega) = 1 \)

iii) for every events \( A \) and \( B \) such that \( A \cap B = \emptyset \) (disjoint events), \( P(A \cup B) = P(A) + P(B) \)
Random Variables

- Again two coins. \( \Omega = \{HH, HT, TH, TT\} \)

- Random Variable \( X : \Omega \to \mathbb{R} \) so that \( X(HH) = 2 \), \( X(HT) = X(TH) = 1 \), \( X(TT) = 0 \), i.e. the number of heads in the outcome.

- Hence, it is sufficient to specify the probabilities \( P(X = x) \) where \( x = 0, 1, 2 \).

- If \( H \) occurs with probability \( p \). Then
  \[
  P(X = 2) = p^2, \quad P(X = 1) = 2p(1-p), \quad P(X = 0) = (1-p)^2.
  \]

- There are many classical distributions. Bernoulli, Binomial, Poisson, Geometrical Distribution, Negative Binomial, etc. ... Normal etc.

- **Why distribution?**
$\Omega = \{H, T\}$. The random variable $X(H) = 1$, $X(T) = 0$.

$\trianglerIGHT P(X = 1) = p$, $P(X = 0) = 1 - p$. Bernoulli.

$\trianglerIGHT$ The Histogram, The Distribution ...

$\trianglerIGHT$ Theoretical vs Real-Life Repeated Outcomes

$\trianglerIGHT$ $X_1, X_2, \ldots, X_n$ - independent Bernoulli($p$)

$\trianglerIGHT$ $Y = X_1 + X_2 + \cdots + X_n$ - Binomial($n, p$)

Flip $n$ coins and count the number of heads

$\trianglerIGHT P(Y = k) = \binom{n}{k} p^k (1 - p)^{n-k}$, where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

http://members.shaw.ca/ron.blond/TLE/Bin.APPLET/index.html
http://zoonek2.free.fr/UNIX/48_R/07.html
Lecture 2

November 15, 2007
Poisson

\[ X = 0, 1, 2, 3, \ldots \]

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

Parameter \( \lambda \), also called the intensity

The probability of a number of events occurring in a fixed period of time if these events occur with a known average rate \( \lambda \)

Examples:

- The number of cars that pass through a certain point on a road during a given period of time
- The number of spelling mistakes one makes while typing a single page
- The number of phone calls at a call center per minute
- The number of times a web server is accessed per minute
- The number of mutations in a given stretch of DNA after a certain amount of radiation
- The number of unstable nuclei that decayed within a given period of time in a piece of radioactive substance

http://en.wikipedia.org/wiki/Poisson_distribution
**Other Discrete Distributions**

**Geometric distribution** \((p)\)

\[ P(X = k) = p(1 - p)^{k-1}, \quad k = 1, 2, \ldots \]

Waiting time until first success in Bernoulli\((p)\) independent trials

**Negative Binomial** \((r, p)\)

\[ P(X_r = k) = \binom{k - 1}{r - 1} p^r (1 - p)^{k-r}, \quad k = r, r + 1, \ldots \]

Waiting time for \(r\)-th success.

HW: Suppose that the number of typographical errors on a single page of a given book has a Poisson distribution with parameters \(\lambda = \frac{1}{2}\). Find the probability that there are at least two errors on page \#10 (the book has more than 10 pages).

HW: An urn contains \(N\) white and \(M\) black balls. Balls are randomly selected, one at a time, with replacement, until a black one is obtained. What is the probability that exactly \(n\) draws are needed.

hint: use Geometric\((p)\). the problem is to guess \(p\).
Suppose $X$ takes all real values $\mathbb{R}$. It does not make any mathematical sense to define $P(X = x)$ for all $x \in \mathbb{R}$.

Specify a function $f_X(x)$, called the density function such that $f \geq 0, \int_{\mathbb{R}} f(x)dx = 1$ and put

$$P(X \in (a, b)) = \int_{a}^{b} f_X(x)dx$$

How does $f$ look like? Almost the histogram (scaled).

**Uniform distribution** $(a, b)$ $f(x) = 1/(b - a)$ for $x \in [a, b]$ and 0 otherwise.

$X \sim \mathcal{N}(\mu, \sigma^2)$, called **Normal Distribution** with mean $\mu$ and standard deviation $\sigma$ if

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

$\mathcal{N}(0, 1)$ - standard normal.
Mean/Expectation and Variance

$X$ - discrete, then the **expectation** or mean is defined

$$
\mathbb{E}(X) = \sum_x x P(X = x).
$$

The average of the possible values of $X$, each value being weighted by its probability.

$X$ - continuous, then $\mathbb{E}(X) = \int f_X(x) dx$.

In general, $\mathbb{E}(g(X)) = \int g(x) f_X(x) dx$.

**Variance.**  
$$
\text{Var}(X) = \mathbb{E}(X - \mathbb{E}(X))^2 = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2.
$$

The amount by which $X$ tends to deviate from its mean.  
$\sigma = \sqrt{\text{Var}(X)}$ is called the **standard deviation**.

HW: Find the mean and variance of Uniform(0,1)
Examples

Bernoulli\((p)\), \(\mathbb{E}(X) = p\), \(\text{Var}(X) = p(1 - p)\)

Binomial\((n, p)\), \(\mathbb{E}(X) = np\), \(\text{Var}(X) = np(1 - p)\)

Geometric\((p)\), \(\mathbb{E}(X) = p^{-1}\), \(\text{Var}(X) = p^{-2}(1 - p)\)

Poisson\((\lambda)\), \(\mathbb{E}(X) = \lambda\), \(\text{Var}(X) = \lambda\)

Negative Binomial \((r, n)\), \(\mathbb{E}(X) = np^{-1}\), \(\text{Var}(X) = np^{-2}(1 - p)\)

\(\mathcal{N}(\mu, \sigma)\), \(\mathbb{E}(X) = \mu\), \(\text{Var}(X) = \sigma^2\)
Central Limit Theorem
Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables, each with mean $\mu$ and variance $\sigma^2$. Then

$$
\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma \sqrt{n}} \to \mathcal{N}(0, 1), \quad n \to \infty.
$$

The convergence is in distribution, i.e.

$$
P\left( \frac{\sum_{i=1}^{n} X_i - \mu}{n} \leq z \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{x^2}{2}} dx
$$

See the histogram.
Run the applet. [http://www.stat.sc.edu/~west/javahtml/CLT.html](http://www.stat.sc.edu/~west/javahtml/CLT.html)
See the PDF’s from Matlab simulations
$Y = \text{Bin}(n, p) = \sum_{i=1}^{n} \text{Bernoulli}(p)$. Hence, $\mu = p$, $\sigma = \sqrt{p(1-p)}$ and

$$\frac{Y - np}{\sqrt{np(1-p)}} \sim \mathcal{N}(0, 1)$$

Note, for $n$ large there is no way to find $P(Y = k)$ directly, while the standard normal is known for all $z$. 
How to find $\mu$ and $\sigma$?

Suppose that $x_1, x_2, \ldots, x_n$ are $n$ realizations of $X$ (a population). Then

$$\hat{\mu} = \frac{\sum_{k=1}^{n} x_k}{n} \approx \mu$$

$\hat{\mu}$ - population mean; $\mu$ the real mean or sample mean. WHY?

**Law of Large Numbers**

If $X_1, \ldots, X_n, \ldots$ i.i.d. Then, $\frac{1}{n} \sum_{i=1}^{n} X_i \rightarrow \mu, \quad n \rightarrow \infty$.

Similarly, the approximation for the standard deviation

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i^2 - \mu) \approx \sigma^2.$$
$X_i$ takes values 1 and $-1$ with probability $p$ and $1 - p$. Independent.

$$S_n = S_0 + \sum_{i=1}^{n} X_i = S_{n-1} + X_n$$

$S_n$ - Random Walk. if $p = 1/2$ - symmetrical random walk.

**Absorbing barrier or Gambler’s ruin**

A Jaguar costs $N$, and a gambler has an initial wealth of $k$, with $0 < k < N$. The gambler plays with a banker the following game: he tosses a coin (could be unfair) repeatedly; if the coin comes up head the banker pays $1, if tail the gambler looses $1. The game ends if the gambler has enough money to buy a Jaguar or he is bankrupted. Find the probability that he is ultimately bankrupted.
It is a random walk!!! Let $p_k$ be the probability of bankruptcy with starting wealth $k$. Then, by conditional probabilities (not very hard to get this)

$$p_k = p \cdot p_{k+1} + (1-p) \cdot p_{k-1} \quad 1 \leq k \leq N - 1$$

with boundary conditions $p_0 = 1$, $p_N = 0$.

It is a difference equation (finite difference), and the solution is

$$ p_k = \left( \frac{1-p}{p} \right)^k - \left( \frac{1-p}{p} \right)^N \cdot \frac{1 - \left( \frac{1-p}{p} \right)^N}{1 - \left( \frac{1-p}{p} \right)^N} .$$

(1)

For $p = 1/2$, i.e. fair coin, we have $p_k = 1 - \frac{k}{N}$.

HW: $p$ in (1) stands for probability of the coin coming up H or T. Which one?

Hint: you do not have to solve the entire problem from the very beginning, just think what this formula means.