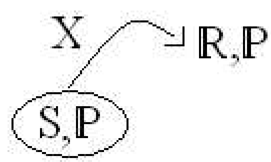
18.05 Lecture 8
February 22, 2005

§3.1 - Random Variables and Distributions
Transforms the outcome of an experiment into a number.
Definitions:
Probability Space: (S, $\mathcal{A}, \mathbb{P}$)
S - sample space, $\mathcal{A}$ - events, $\mathbb{P}$ - probability
Random variable is a function on S with values in real numbers, X:S $\rightarrow \mathbb{R}$

Examples:
Toss a coin 10 times, Sample Space = {HTH...HT, ...}, all configurations of H & T.
Random Variable X = number of heads, X: S $\rightarrow \mathbb{R}$
X: S $\rightarrow \{0, 1, ..., 10\}$ for this example.
There are fewer outcomes than in S, you need to give the distribution of the random variable in order to get the entire picture. Probabilities are therefore given.

Definition: The distribution of a random variable X:S $\rightarrow \mathbb{R}$, is defined by: $A \subseteq \mathbb{R}, \mathbb{P}(A) = \mathbb{P}(X \in A)$ $= \mathbb{P}(s \in S : X(s) \in A)$



The random variable maps outcomes and probabilities to real numbers.
This simplifies the problem, as you only need to define the mapped $\mathbb{R}, \mathbb{P}$, not the original S, $\mathbb{P}$.
The mapped variables describe X, so you don't need to consider the original complicated probability space.

From the example, $\mathbb{P}(X = \#(\text{heads in 10 tosses}) = k) = \binom{10}{k}(\frac{1}{2})^k(\frac{1}{2})^{10-k} = \binom{10}{k}\frac{1}{2^{10}}$
Note: need to distribute the heads among the tosses,
account for probability of both heads and tails tossed.
This is a specific example of the more general binomial problem:
A random variable X $\in \{1, ..., n\}$

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

This distribution is called the **binomial distribution**: B(n, p),
which is an example of a discrete distribution.

Discrete Distribution
A random variable X is called discrete if it takes a finite or countable number (sequence) of values: $X \in \{s_1, s_2, s_3, ...\}$
It is completely described by telling the probability of each outcome.
Distribution defined by: $\mathbb{P}(X = s_k) = f(s_k)$, the probability function (p.f.)
p.f. cannot be negative and should sum to 1 over all outcomes.
$\mathbb{P}(X \in A) = \sum_{s_k \in A} f(s_k)$

Example: Uniform distribution of a finite number of values {1, 2, 3, ..., n} each outcome

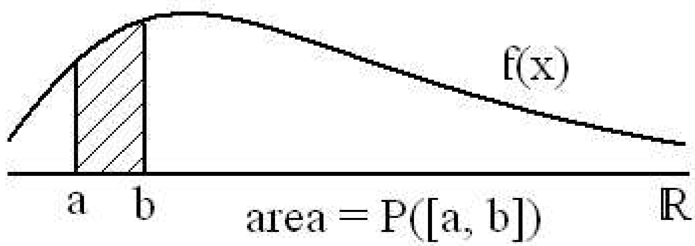
has equal probability $\rightarrow f(s_k) = \frac{1}{n}$: uniform probability function.
random variable X $\in \mathbb{R}, \mathbb{P}(A) = \mathbb{P}(X \in A), A \subseteq \mathbb{R}$
can redefine probability space on random variable distribution:
$(\mathbb{R}, \mathcal{A}, \mathbb{P})$ - sample space, X: $\mathbb{R} \rightarrow \mathbb{R}, X(x) = x$ (identity map)
$\mathbb{P}(A) = \mathbb{P}(X : X(x) \in A) = \mathbb{P}(x \in A) = \mathbb{P}(x \in A) = \mathbb{P}(A)$
all you need is the outcomes mapped to real numbers and relative probabilities of the mapped outcomes.

Example: **Poisson Distribution**, {0, 1, 2, 3, ...} $\Pi(\lambda), \lambda =$ intensity
probability function:

$$f(k) = \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \text{ where } \lambda \text{ parameter } > 0.$$

$\sum_{k \geq 0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k \geq 0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = e^0 = 1$
Very common distribution, will be used later in statistics.
Represents a variety of situations - ex. distribution of "typos" in a book on a particular page, number of stars in a random spot in the sky, etc.
Good approximation for real world problems, as $\mathbb{P} > 10$ is small.

Continuous Distribution
Need to consider intervals not points.
Probability distribution function (p.d.f.): $f(x) \geq 0$.
Summation replaced by integral: $\int_{-\infty}^{\infty} f(x)dx = 1$
then, $\mathbb{P}(A) = \int_A f(x)dx$, as shown:



If you were to choose a random point on an interval, the probability of choosing a particular point is equal to zero.
You can't assign positive probability to any point, as it would add up infinitely on a continuous interval.
It is necessary to take P(point is in a particular sub-interval).
Definition implies that $\mathbb{P}(\{a\}) = \int_a^a f(x)dx = 0$

Example: In a uniform distribution [a, b], denoted U[a, b]:
p.d.f.: $f(x) = \frac{1}{b-a}$, for x $\in$ [a, b]; 0, for x $\notin$ [a, b]

Example: On an interval [a, b], such that $a < c < d < b$,
$\mathbb{P}([c, d]) = \int_c^d \frac{1}{b-a} dx = \frac{d-c}{b-a}$ (probability on a subinterval)

Example: **Exponential Distribution**

$$E(\alpha), \alpha > 0 \text{ parameter}$$

p.d.f.: $f(x) = \alpha e^{-\alpha x}, \text{ if } x \geq 0; 0, \text{ if } x < 0$
Check that it integrates to 1:

$\int_0^\infty \alpha e^{-\alpha x} dx = \alpha(-\frac{1}{\alpha} e^{-\alpha x}|_0^\infty = 1$
Real world: Exponential distribution describes the life span of quality products (electronics).

** End of Lecture 8