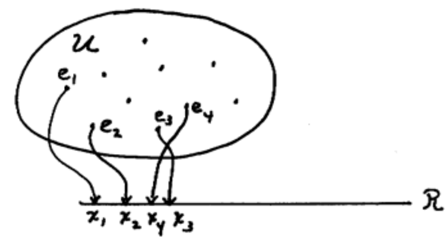
Random Variables

Lecture 2

Spring Quarter, 2002

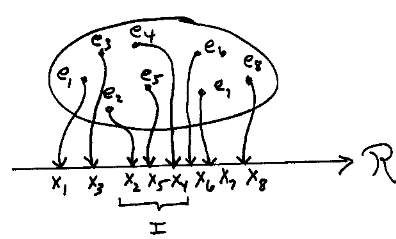
Definition

Let $\mathbf{E}$ be an experiment whose outcomes are a sample space $\mathcal{U}$ for which the probability $P(e)$ is defined for each outcome $e \in \mathcal{U}$. A random variable is a function $X(e)$ that associates a number with each outcome $e \in \mathcal{U}$.



Intervals

- Every interval on $\mathcal{R}$ corresponds to a set of outcomes in $\mathcal{U}$.
- Let $I \in \mathcal{R}$ be an interval. Then $\mathcal{A} = \{e \in \mathcal{U} : X(e) = x, x \in I\}$ is an event.
- $P(\mathcal{A})$ can be calculated, and $P(I) = P(\mathcal{A})$
- Every interval of R is an event whose probability can be calculated.
- For the figure below, $\mathcal{A} = \{e_2, e_4, e_5\}$ and $P(I) = P(\mathcal{A}) = P(e_2) + P(e_4) + P(e_5)$



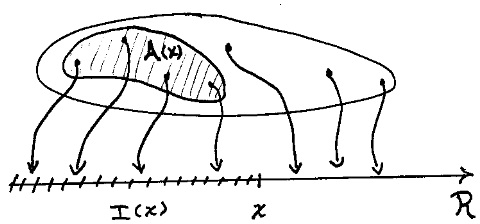
Probability Distribution Function

- Consider the semi-infinite interval $I_x = \{s : s \le x\}$ be the interval to the left of x.
- Let $X(e)$ be a random variable.
- Let $\mathcal{A}(x)$ be the event that $X \in I(x)$ Then $\mathcal{A}(x) = \{e : X(e) \le x\}$.
- The probability $P(X \in I) = P(X \le x) = P(\mathcal{A}(x))$ is well defined for every x.

The probability $P(X \le x)$ is a special function of x called the *probability distribution function.*

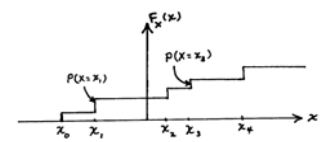
$$F_X(x) = P(X \le x)$$

Probability Distribution Function



$$\lim_{x \to -\infty} F_X(x) = 0$$
$$\lim_{x \to \infty} F_X(x) = 1$$
$$P(a < x \leq b) = F_X(b) - F_X(a)$$

Discrete Distribution



A discrete distribution function has a finite number of discontinuities. The random variable has a nonzero probability only at the points of discontinuity.

The distribution function for a discrete random variable is a staircase that increases from left to right.

Continuous Distribution

Suppose that $F_X(x)$ is continuous for all x. Then

$$\lim_{\varepsilon \to 0} F_X(x) - F_X(x - \varepsilon) = 0$$

so that $P(X = x) = 0$ for all x.

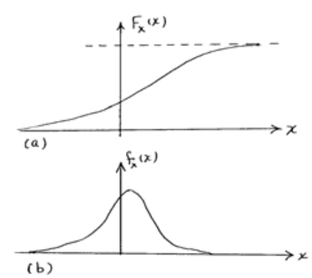
The derivative is well-defined where $F_X(x)$ is continuous.

$$\frac{dF_X(x)}{dx} = \lim_{\varepsilon \to 0} \frac{F_X(x) - F_X(x - \varepsilon)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{P(x - \varepsilon < X \leq x)}{\varepsilon}$$

The slope of the probability distribution function is equivalent to the *density* of probability.

$$f_X(x) = \frac{dF_X(x)}{dx}$$

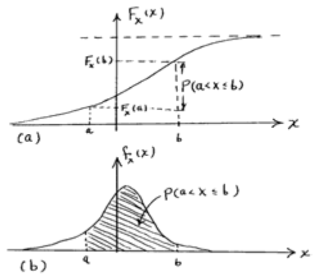
Continuous Distribution



The distribution function (a) for a continuous random variable and (b) its probability density function. Note that the probability density function is highest where the slope of the distribution function is greatest.

Continuous Distribution

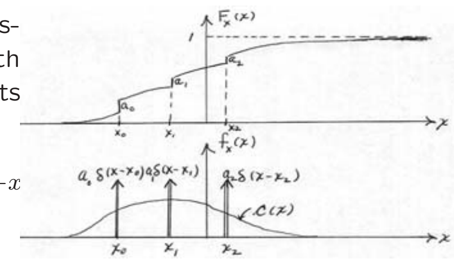
$$f_X(x) = \frac{dF_X(x)}{dx}$$
$$F_X(x) = \int_{-\infty}^{x} f_X(u)du$$
$$P(a < X \le b) = F_X(b) - F_X(a) = \int_{a}^{b} f_X(u)du$$



The probability $P(a < X \le b)$ is related to the change in height of the distribution and to the area shown in the probability density function.

Mixed Distribution

The range of a mixed distribution contains isolated points and points in a continuum. The distribution function is a smooth curve except at one or more points where there are finite steps.



$$f_X(x) = c(x) + \sum_k P(X = x_k)\delta(x - x$$

$$c(x) = dF_X/dx$$

where $F(x)$ is continuous.

Random Vector

Let $\mathbf{E}$ be an experiment whose outcomes are a sample space $\mathcal{U}$ for which the probability $P(e)$ is defined for each outcome $e \in \mathcal{U}$.

A random vector is a function $\mathbf{X}(e) = [X_1(e), X_2(e), \ldots, X_r(e)]$ where $X_i(e)$ $i = 1, 2, \ldots, r$ are random variables defined over the space $\mathcal{U}$.

Joint Probability Distribution Function

$$F_{X_1X_2}(x_1, x_2) = P[(X_1 \le x_1) \cap (X_2 \le x_2$$

1. $F_{X_1X_2}(-\infty, -\infty) = 0$
2. $F_{X_1X_2}(-\infty, x_2) = 0$ for any x_2
3. $F_{X_1X_2}(x_1, -\infty) = 0$ for any x_1
4. $F_{X_1X_2}(+\infty, +\infty) = 1$
5. $F_{X_1X_2}(+\infty, x_2) = F_{X_2}(x_2)$ for any x_2
6. $F_{X_1X_2}(x_1, +\infty) = F_{X_1}(x_1)$ for any x_1

Joint Probability Distribution Function

The probability that an experiment produces a pair (X_1, X_2) that falls in a rectangular region with lower left corner (a, c) and upper right corner (b, d) is

$$
\begin{aligned}
P[(a < X_1 \leq b) \cap (c < X_2 \leq d)] \; = \; & F_{X_1X_2}(b,d) - F_{X_1X_2}(a,d) - \\
& F_{X_1X_2}(b,c) + F_{X_1X_2}(a,c)
\end{aligned}
$$

Joint Probability Density Function

$$f_{X_1X_2}(x_1, x_2) = \frac{\partial^2 F_{X_1X_2}(x_1, x_2)}{\partial x_1 \partial x_2}$$

$$
\begin{aligned}
f_{U,V}(u,v) &\geq 0 \\
F_{U,V}(u,v) &= \int_{-\infty}^{u} \int_{-\infty}^{v} f_{U,V}(\xi,\eta) d\xi d\eta \\
\int_{-\infty}^{\infty} f_{U,V}(\xi,\eta) d\xi d\eta &= 1 \\
F_U(u) &= \int_{-\infty}^{u} \int_{-\infty}^{\infty} f_{U,V}(\xi,\eta) d\xi d\eta \\
F_V(v) &= \int_{-\infty}^{\infty} \int_{-\infty}^{v} f_{U,V}(\xi,\eta) d\xi d\eta \\
f_U(u) &= \int_{-\infty}^{\infty} f_{U,V}(u,\eta) d\eta \\
f_V(v) &= \int_{-\infty}^{\infty} f_{U,V}(\xi,v) d\xi
\end{aligned}
$$

Die Tossing Example

Mapping of the outcomes of the die tossing experiment onto points in a plane by a particular pair of random variables.

Each outcome maps into a pair of random variables.