Chapter 2 in Miller and Childers

We are carrying out a <u>random experiment</u>. greek xi

The experiment has various possible <u>outcomes</u> ξ.

The <u>sample space</u> S is the set of all possible outcomes.

An <u>event</u> A is a subset of the sample space $A \subseteq S$.

A <u>probability (measure)</u> Pr assigns probabilities $Pr(A)$ to all events $A \subseteq S$.

<u>Axioms for Probability Measures</u>

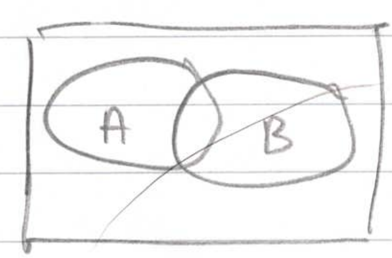
(1) $Pr(A) \geq 0$ for $A \subseteq S$

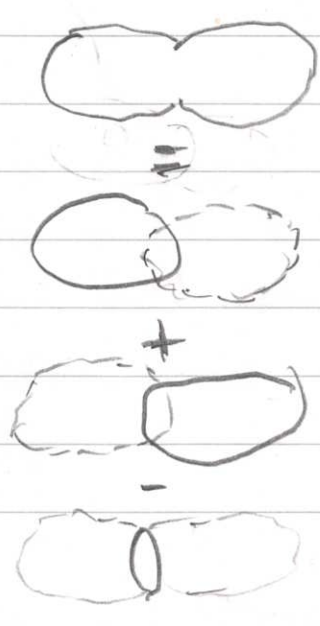
(2) $Pr(S) = 1$

(3) $Pr(A \cup B) = Pr(A) + Pr(B)$ for $A, B \subseteq S$ with $A \cap B = \emptyset$

<u>Theorem</u> $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$ for $A, B \subseteq S$.

<u>Proof</u>





#

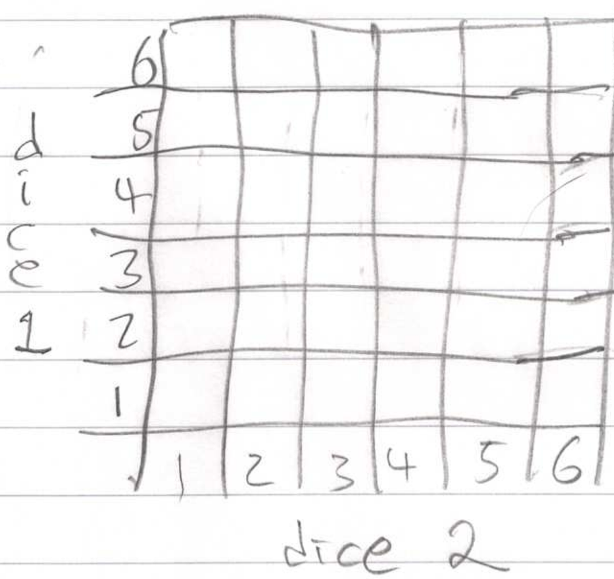
Theorem $\Pr(\bar{A}) = \Pr(A^c) = 1 - P(A)$

complement to $A = S \setminus A$

Proof $1 = \Pr(S) = \Pr(A \cup A^c) = \Pr(A) + \Pr(A^c)$. #

since $A \cap A^c = \emptyset$

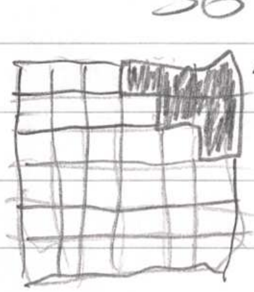
Example 1 Toss with two dice

$S =$ 

$= S$

$\Pr(\text{each } \square) = \frac{1}{36}$

$$\Pr(A) = \sum_{\square \in A} \Pr(\square) = \frac{\#\{\square \in S : \square \in A\}}{36}$$

$A = \{\text{sum of dice} \geq 10\} =$  6 slots of 36

$\Pr(A) = \frac{6}{36} = \frac{1}{6}$

Joint probabilities $\Pr(A, B) = \Pr(A \cap B)$

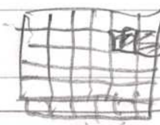
book math

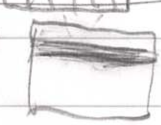
two different notations for same

Definition The conditional probability for $A \subseteq S$ given that $B \subseteq S$ occures (occured) $Pr(A|B)$ is (defined as) $\frac{Pr(A \cap B)}{Pr(B)}$

Example 1 (continued) If B is first dice shows 5 then

$$Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)} = \frac{Pr(\quad)}{Pr(\quad)}$$





$$= \frac{2/36}{1/6} = \frac{2}{6}$$

Theorem of Total Probability Let $B_1, \ldots, B_n$ be mutually exclusive (nonoverlapping, i.e., $B_i \cap B_j = \phi$ for $i \neq j$) events with $\bigcup_{i=1}^{n} B_i = S$. Then

$$Pr(A) = \sum_{i=1}^{n} Pr(A|B_i)\, Pr(B_i) \quad \text{for } A \subseteq S$$

Proof $\sum_{i=1}^{n} Pr(A|B_i)\, Pr(B_i) = \sum_{i=1}^{n} \frac{Pr(A \cap B_i)}{Pr(B_i)} Pr(B_i)$

$$= \sum_{i=1}^{n} Pr(A \cap B_i) = Pr\left(\bigcup_{i=1}^{n} (A \cap B_i)\right) = Pr(A) \quad //$$

Bayes' Theorem For $B_1, \ldots, B_n$ as in the previous theorem we have

$$P(B_i|A) = \frac{Pr(A|B_i)\, Pr(B_i)}{\sum_{i=1}^{n} Pr(A|B_i)\, Pr(B_i)}$$

Proof Numerator is $Pr(A \cap B_i)$ and denominator $Pr(A)$. $\#$

Definition Two events $A, B \subseteq S$ are independent if $Pr(A \cap B) = Pr(A)\,Pr(B)$.

Consequence $Pr(A|B) = Pr(A)$ for $A, B \subseteq S$ independent.

Definition n events $A_1, \dots, A_n \subseteq S$ are independent if $Pr(A_{i_1} \cap \dots A_{i_k}) = Pr(A_{i_1}) \cdot \dots \cdot Pr(A_{i_k})$ for any choice of distinct $i_1, \dots, i_k \in \{1, \dots, n\}$.

Example 1 (continued) Are $A = \{\text{sum of dice} \geq 10\}$ and $B = \{\text{dice one shows } 5\}$ independent?

Answer: As $Pr(A) = \frac{1}{6}$, $Pr(A \cap B) = \frac{2}{36}$ and $Pr(B) = \frac{1}{6}$ they are (as is quite clear from common sense) NOT!

Definition A random variable (r.v.) $X = X(s)$ is a function $X : S \to \mathbb{R}$ from a sample space to the real numbers.

Definition A random variable is discrete if the number of possible values of it is finite or countably infinite. If the number of possible values instead is uncountably infinite the random variable is called continuous.

Definition The Probability Mass Function (PMF) p_X is defined $p_X(x) = Pr(X = x)$ for discrete X.

(The theorem last on page 2.6 should come here.)

Example 1 (continued) If X is sum of dice

$P_X(x) = \frac{1}{36}, \frac{2}{36}, \frac{3}{36}, \frac{4}{36}, \frac{5}{36}, \frac{6}{36}, \frac{5}{36}, \frac{4}{36}, \frac{3}{36}, \frac{2}{36}, \frac{1}{36}$ for $x = 2, \dots, 12$ respectively.

Bernoulli r.v. X has possible values $\{0,1\}$ with $P_X(0) = 1-p$ and $P_X(1) = p$ for a $p \in [0,1]$.

X can be viewed as indicator if coin toss is ~~tails or~~ heads.

Binomial r.v. X has possible values $\{0, \dots, n\}$ with $P_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$ for $x = 0, \dots, n$.

X can be viewed as the number of heads obtained in a total of n independent coin tosses since the probability of a certain ordered sequence of x heads and $n-x$ tails has probability $p^x(1-p)^{n-x}$ and there are $\binom{n}{x}$ different such ordered sequences.

Poisson r.v. X has possible values $\mathbb{N} = \{0, 1, \dots\}$ with $P_X(x) = \frac{\alpha^x}{x!} e^{-\alpha}$ for $x = 0, 1, 2, \dots$ for some parameter $\alpha > 0$.

Poisson r.v. is often used to model number of arrivals during a certain time, e.g., number of cars passing on a road during one hour.

Waiting time (geometric) r.v. X

has possible values $x \in \{1, 2, \ldots\}$ with $P_X(x) = (1-p)\,p^{x-1}$ which is a model for the number X of independen coin tosses needed to obtain first tails (meaning first $x-1$ heads followed by 1 tails).

Theorem For a PMF we have

(1) $P_X(x) \geq 0$

(2) $\Pr(X \in A) = \sum_{x \in A} P_X(x)$ for $A \subseteq \mathbb{R}$

(3) $\sum_{\text{all } x} P_X(x) = 1$.