Probability

Suppose that we have such an experiment, but the experiment is of such a nature that a collection of every possible outcome can be described prior to its performance.
If this kind of experiment can be repeated under the same conditions, it is called a **random experiment**, and the collection of every possible outcome is called the experimental space or the **sample space**. We denote the sample space by C.

Example 1.1.1. In the toss of a coin, let the outcome tails be denoted by T and let the outcome heads be denoted by H. If we assume that the coin may be repeatedly tossed under the same conditions, then the toss of this coin is an example of a random experiment in which the outcome is one of the two symbols T or H; that is, the sample space is the collection of these two symbols. For this example, then, C = {H, T}.

Example 1.1.2. In the cast of one red die and one white die, let the outcome be the ordered pair (number of spots up on the red die, number of spots up on the white die). If we assume that these two dice may be repeatedly cast under the same conditions, then the cast of this pair of dice is a random experiment. The sample space consists of the 36 ordered pairs:
C = {(1, 1), . . . , (1, 6), (2, 1), . . . , (2, 6), . . . , (6, 6)}.

Subsets of C are often called **events** and are generally denoted by capitol Roman letters such as A,B, or C.

If the experiment results in an element in an event A, we say the event A has occurred. We are interested in the chances that an event occurs.

For instance, in Example 1.1.1 we may be interested in the chances of getting heads; i.e., the chances of the event A = {H} occurring.
In the second example, we may be interested in the occurrence of the sum of the upfaces of the dice being “7” or “11;” that is, in the occurrence of the event A = {(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1), (5, 6), (6, 5)}

We assume that in all cases, the collection of events is sufficiently rich to include all possible events of interest and is closed under complements and countable unions of these events. Using DeMorgan’s Laws, the collection is then also closed under countable intersections.

We denote this collection of events by B. Technically, such a collection of events is called a **σ-field** of subsets.

\

Probability set function
Now that we have a sample space, C, and our collection of events, B, we can define the third component in our probability space, namely a probability set function

Definition 1.3.1 (Probability)**.** Let C be a sample space and let B be the set of events. Let P be a real-valued function defined on B. Then P is a **probability set function** if P satisfies the following three conditions:
1. P(A) ≥ 0, for all A ∈ B.
2. P(C) = 1.
3. If {An} is a sequence of events in B and Am ∩ An = φ for all m = n, then

$$P\left(\cup_{n=1}^{\propto} A_n\right) = \sum_{n=1}^{\propto} P(An)$$

A collection of events whose members are pairwise disjoint, as in (3), is said to be a **mutually exclusive** collection and its union is often referred to as a **disjoint union**. The collection is further said to be **exhaustive** if the union of its events is the sample space, in which case
$\sum_{n=1}^{\propto} P(An) = 1$.
We often say that a mutually exclusive and exhaustive collection of events forms a **partition** of C.

Theorem 1.3.1. For each event A ∈ B, P(A) = 1 − P(Ac).
Proof: We have C = A ∪ Ac and A ∩ Ac = φ. Thus, from (2) and (3) of Definition 1.3.1, it follows that
1 = P(A) + P(Ac),
which is the desired result.

Theorem 1.3.2. The probability of the null set is zero; that is, P(φ) = 0.
Proof: In Theorem 1.3.1, take A = φ so that Ac = C. Accordingly, we have
P(φ) = 1 − P(C) = 1 −1 = 0
and the theorem is proved.

Theorem 1.3.3. If A and B are events such that A ⊂ B, then P(A) ≤ P(B).
Proof: Now B = A∪ (Ac ∩B) and A∩ (Ac ∩B) = φ. Hence, from (3) of Definition 1.3.1,
P(B) = P(A) + P(Ac ∩ B).
From (1) of Definition 1.3.1, P(Ac ∩ B) ≥ 0. Hence, P(B) ≥ P(A).

Theorem 1.3.4. For each A ∈ B, 0 ≤ P(A) ≤ 1.
Proof: Since φ ⊂ A ⊂ C, we have by Theorem 1.3.3 that
P(φ) ≤ P(A) ≤ P(C) or 0≤ P(A) ≤ 1,
the desired result.

Theorem 1.3.5. If A and B are events in C, then
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
Proof: Each of the sets $A \cup B$ and B can be represented, respectively, as a union of nonintersecting sets as follows:

$A \cup B = A \cup (Ac \cap B)$ and $B = (A \cap B) \cup (Ac \cap B)$.

That these identities hold for all sets A and B follows from set theory.

Thus, from (3) of Definition 1.3.1,
$P(A \cup B) = P(A) + P(Ac \cap B)$
and
$P(B) = P(A \cap B) + P(Ac \cap B)$.

If the second of these equations is solved for $P(Ac \cap B)$ and this result is substituted in the first equation, we obtain

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
This completes the proof.

Example 1.3.1. Let C denote the sample space of Example 1.1.2. Let the probability set function assign a probability of 1/36 to each of the 36 points in C; that is, the dice are fair. If C1 = {(1, 1), (2, 1), (3, 1), (4, 1), (5, 1)} and C2 = {(1, 2), (2, 2), (3, 2)}, then P(C1) = 5/36 , P(C2) = 3/36 , $P(C1 \cup C2) = 8/36$, and $P(C1 \cap C2) = 0$.

Example 1.3.2. Two coins are to be tossed and the outcome is the ordered pair (face on the first coin, face on the second coin). Thus the sample space may be represented as C = {(H,H), (H, T), (T,H), (T,T)}. Let the probability set function assign a probability of 1/4 to each element of C. Let C1 = {(H,H), (H, T)} and C2 = {(H,H), (T,H)}. Then P(C1) = P(C2) = 1/2 , $P(C1 \cap C2) = 1/4$, and,

In accordance with Theorem 1.3.5, $P(C1 \cup C2) = 1/2 + \frac{1}{2} - 1/4 = 3/4$.

Definition 1.3.2 (Equilikely Case)**.** Let C = {x1, x2, . . . , xm} be a finite sample space. Let pi = 1/m for all i = 1, 2, . . .,m and for all subsets A of C define
$P(A) = \sum_{xi \in A} \frac{1}{m} = \frac{n(A)}{m}$,
where n(A) denotes the number of elements in A. Then P is a probability on C and it is refereed to as the **equilikely case**.

Example 1.3.4 (Poker Hands)**.**

Let a card be drawn at random from an ordinary deck of 52 playing cards that has been well shuffled.

a) If , if E1 is the set of outcomes that are spades, and E2 is the set of outcomes that are kings, Find P(E1) and P(E2)

Solution (a):

The sample space C consists of 52 elements, each element represents one and only one of the 52 cards.

Because the deck has been well shuffled, it is reasonable to assume that each of these outcomes has the same probability 1/52 .

Accordingly, if E1 is the set of outcomes that are spades, P(E1) = 13/52 = 1/4 ,
because there are 13 spades in the deck; that is, 1/4 is the probability of drawing a card that is a spade.

If E2 is the set of outcomes that are kings, P(E2) = 4/52 = 1/13
, because there are 4 kings in the deck; that is, 1/13 is the probability of drawing a card that is a king.

(b) Now, suppose that five cards are taken, at random and without replacement, from this deck; i.e, a five card poker hand. In this instance, order is not important. So a hand is a subset of five elements drawn from a set of 52 elements. Hence, there are ${}^{52}C_5$ poker hands.

If the deck is well shuffled, each hand should be equilikely; i.e., each hand has probability $1/{}^{52}C_5$.

We can now compute the probabilities of some interesting poker hands.

Let E1 be the event of a flush, all five cards of the same suit. There are ${}^{4}C_1 = 4$ suits to choose for the flush and in each suit there are ${}^{13}C_5$ possible hands; hence, using the multiplication rule, the probability of getting a flush is

$$P(E1) = \frac{4C1 \text{ x } 13C5}{52C5} = \frac{4 \cdot 1287}{2598960} = 0.00198.$$

(c) Next, consider the probability of the event E2 of getting exactly three of a kind, (the other two cards are distinct and are of different kinds).

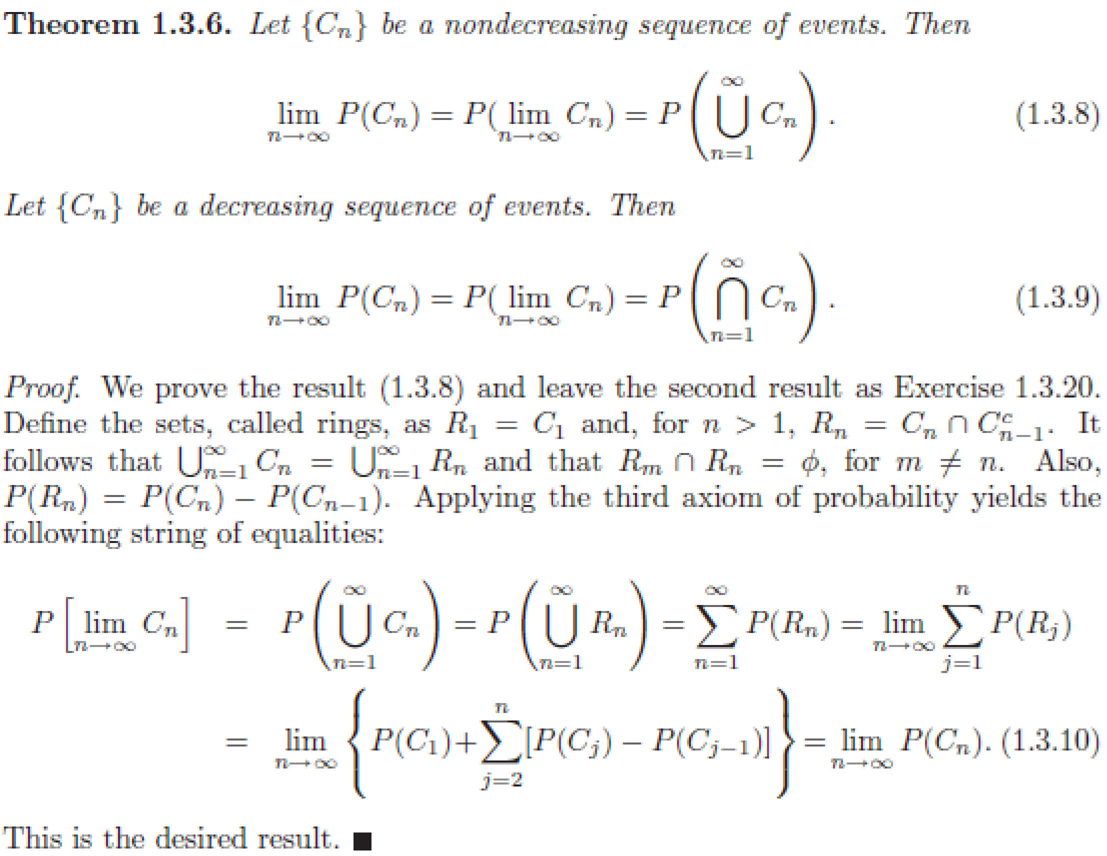
Choose the kind for the three, in ${}^{13}C_1$ ways;

choose the three, in ${}^{4}C_3$ ways; choose the other two kinds, in ${}^{12}C_2$ ways; and choose one card from each of these last two kinds, in ${}^{4}C_1\,{}^{4}C_1$ ways.

Hence the probability of exactly three of a kind is

P(E2) = $\frac{13C1\ \ 4C3\ \ 12C2\ \ 4C1\ 4C1}{52C5}$ = 0.0211.

Definition: A sequence of events {Cn} is a nondecreasing sequence if Cn ⊂ Cn+1, for all n, in which case we wrote limn→∞ Cn = $\cup^{\infty}_{n=1}$Cn.



Theorem 1.3.6. *Let $\{C_n\}$ be a nondecreasing sequence of events. Then*

$$\lim_{n\to\infty} P(C_n) = P(\lim_{n\to\infty} C_n) = P\left(\bigcup_{n=1}^{\infty} C_n\right). \tag{1.3.8}$$

Let $\{C_n\}$ be a decreasing sequence of events. Then

$$\lim_{n\to\infty} P(C_n) = P(\lim_{n\to\infty} C_n) = P\left(\bigcap_{n=1}^{\infty} C_n\right). \tag{1.3.9}$$

Proof. We prove the result (1.3.8) and leave the second result as Exercise 1.3.20. Define the sets, called rings, as $R_1 = C_1$ and, for $n > 1$, $R_n = C_n \cap C_{n-1}^c$. It follows that $\bigcup_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} R_n$ and that $R_m \cap R_n = \phi$, for $m \neq n$. Also, $P(R_n) = P(C_n) - P(C_{n-1})$. Applying the third axiom of probability yields the following string of equalities:

$$\begin{aligned} P\left[\lim_{n\to\infty} C_n\right] &= P\left(\bigcup_{n=1}^{\infty} C_n\right) = P\left(\bigcup_{n=1}^{\infty} R_n\right) = \sum_{n=1}^{\infty} P(R_n) = \lim_{n\to\infty} \sum_{j=1}^{n} P(R_j) \\ &= \lim_{n\to\infty} \left\{ P(C_1) + \sum_{j=2}^{n} [P(C_j) - P(C_{j-1})] \right\} = \lim_{n\to\infty} P(C_n). \quad (1.3.10) \end{aligned}$$

This is the desired result. ■

Theorem 1.3.7 (Boole's Inequality). *Let $\{C_n\}$ be an arbitrary sequence of events. Then*

$$P\left(\bigcup_{n=1}^{\infty} C_n\right) \le \sum_{n=1}^{\infty} P(C_n). \tag{1.3.11}$$

Proof: Let $D_n = \bigcup_{i=1}^{n} C_i$. Then $\{D_n\}$ is an increasing sequence of events that go up to $\bigcup_{n=1}^{\infty} C_n$. Also, for all j, $D_j = D_{j-1} \cup C_j$. Hence, by Theorem 1.3.5,

$$P(D_j) \le P(D_{j-1}) + P(C_j),$$

that is,

$$P(D_j) - P(D_{j-1}) \le P(C_j).$$

In this case, the C_is are replaced by the D_is in expression (1.3.10). Hence, using the above inequality in this expression and the fact that $P(C_1) = P(D_1)$, we have

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} C_n\right) &= P\left(\bigcup_{n=1}^{\infty} D_n\right) = \lim_{n\to\infty} \left\{ P(D_1) + \sum_{j=2}^{n} [P(D_j) - P(D_{j-1})] \right\} \\ &\le \lim_{n\to\infty} \sum_{j=1}^{n} P(C_j) = \sum_{n=1}^{\infty} P(C_n). \quad \blacksquare \end{aligned}$$

Exercise

1.3.6. If the sample space is $\mathcal{C} = \{c : -\infty < c < \infty\}$ and if $C \subset \mathcal{C}$ is a set for which the integral $\int_C e^{-|x|}\, dx$ exists, show that this set function is not a probability set function. What constant do we multiply the integrand by to make it a probability set function?

Solution. With $C = \mathcal{C} = \mathbb{R}$ we have

$$\begin{aligned} \int_{\mathbb{R}} e^{-|x|}\, dx &= \int_{-\infty}^{\infty} e^{-|x|}\, dx \\ &= 2\int_0^{\infty} e^{-|x|}\, dx \text{ since } e^{-|x|} \text{ is an even function} \\ &= 2\int_0^{\infty} e^{-x}\, dx \text{ since } x \ge 0 \text{ here} \\ &= 2\lim_{b\to\infty}\left(\int_0^b e^{-x}\, dx\right) = 2\lim_{b\to\infty}\left(-e^{-x}|_0^b\right) \\ &= 2\lim_{b\to\infty}(-e^{-b}+1) = 2(0+1) = 2. \end{aligned}$$

So $\int_C e^{-|x|}\,dx$ is not a probability set function

If we define $P(C) = \frac{1}{2}\int_C e^{-|x|}\,dx$ then we have

$P(C) = 1$. and Definition 1.3.1(2) is then satisfied.

Take $P(\varnothing) = \int_\varnothing e^{-|x|}\,dx = 0$, so that Definition 1.3.1(1) is satisfied.

Exercise:

1.3.8. Let C_1, C_2, and C_3 be three mutually disjoint subsets of the sample space $\mathcal{C}$. Find $P[(C_1 \cup C_2) \cap C_3]$ and $P(C_1^c \cup C_2^c)$.

Exercise 1.3.9. Determine the probability of being dealt a full house, i.e., three-of-a-kind and two-of-a-kind.

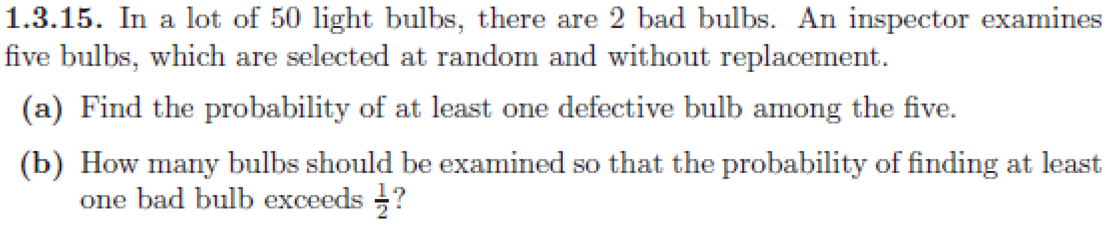
Solution. The suit of the three-of-a-kind can be chosen in $\binom{13}{1} = 13$ ways and the suit of the two-of-a-kind can then be chose in $\binom{12}{1} = 12$ ways. The three cards in the three-of-a-kind can then be chosen in $\binom{4}{3}$ ways and the two cards in the two-of-a-kind can then be chosen in $\binom{4}{2}$ ways. So the probability of being dealt a full house is

$$\frac{\binom{13}{1}\binom{12}{1}\binom{4}{3}\binom{4}{2}}{\binom{52}{5}} = \frac{(13)(12)(4)(6)}{2{,}598{,}960} \approx 0.00144.$$

Exercise

1.3.14. There are five red chips and three blue chips in a bowl. The red chips are numbered 1, 2, 3, 4, 5, respectively, and the blue chips are numbered 1, 2, 3, respectively. If two chips are to be drawn at random and without replacement, find the probability that these chips have either the same number or the same color.

Exercise

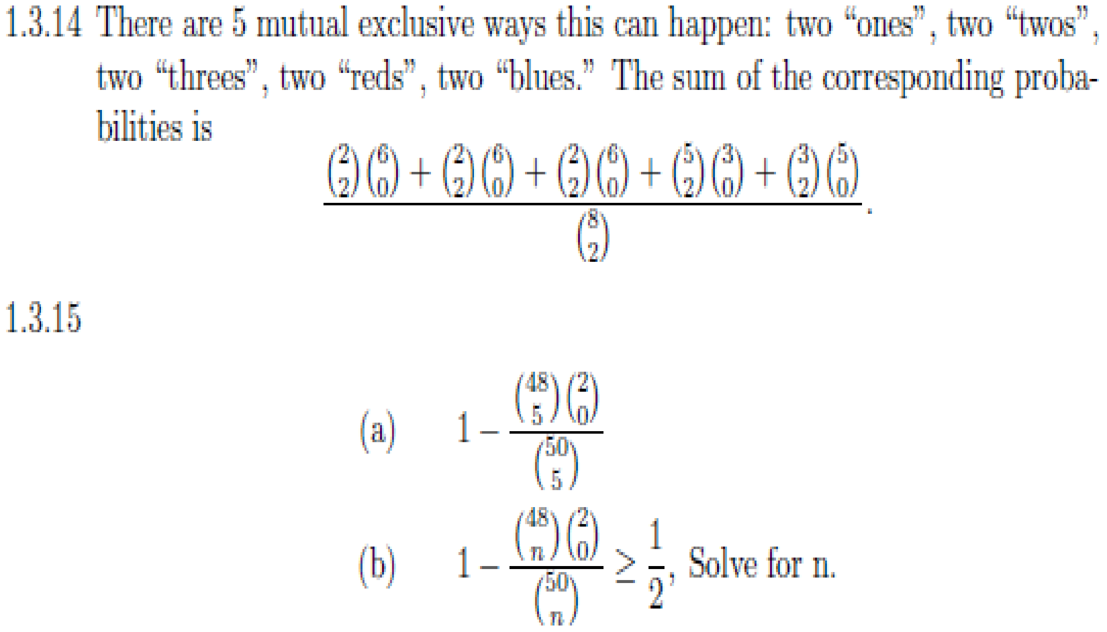


1.3.15. In a lot of 50 light bulbs, there are 2 bad bulbs. An inspector examines five bulbs, which are selected at random and without replacement.

(a) Find the probability of at least one defective bulb among the five.

(b) How many bulbs should be examined so that the probability of finding at least one bad bulb exceeds $\frac{1}{2}$?

Solutions



1.3.14 There are 5 mutual exclusive ways this can happen: two "ones", two "twos", two "threes", two "reds", two "blues." The sum of the corresponding probabilities is

$$\frac{\binom{2}{2}\binom{6}{0}+\binom{2}{2}\binom{6}{0}+\binom{2}{2}\binom{6}{0}+\binom{5}{2}\binom{3}{0}+\binom{3}{2}\binom{5}{0}}{\binom{8}{2}}.$$

1.3.15

$$\text{(a)} \qquad 1-\frac{\binom{48}{5}\binom{2}{0}}{\binom{50}{5}}$$

$$\text{(b)} \qquad 1-\frac{\binom{48}{n}\binom{2}{0}}{\binom{50}{n}} \geq \frac{1}{2}, \text{ Solve for n.}$$
