binomial theorem

probability of the complement

probability of the union of two events

the multiplication rule

Bayes' formula
\[ n \text{ choose } k \text{ is a brief way of saying how many ways can you choose } k \text{ objects from a set of } n \text{ objects, when the order of selection is not relevant.} \]
\[
\binom{n}{k} = \frac{n!}{(n-k)! \cdot k!}
\]

Obviously, this implies \( 0 \leq k \leq n \).

Suppose you want to divide \( n \) distinct items into \( r \) distinct groups each with size \( n_1, n_2, \ldots, n_r \), how do you count the possible outcomes?

If \( n_1 + n_2 + \ldots + n_r = n \), then the number of possible divisions can be counted by the following formula:

\[
\binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1! \cdot n_2! \ldots n_r!}
\]

If \( E^c \) denotes the complement of event \( E \), then
\[
P(E^c) = 1 - P(E)
\]

If \( P(F) > 0 \), then
\[
P(E \mid F) = \frac{P(EF)}{P(F)}
\]

These flashcards and the accompanying \LaTeX{} source code are licensed under a Creative Commons Attribution–NonCommercial–ShareAlike 2.5 License. For more information, see creativecommons.org. You can contact the author at:

jasonu [remove-this] at physics dot utah dot edu

\[
(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}
\]

1. \( 0 \leq P(E) \leq 1 \)
2. \( P(S) = 1 \)
3. For any sequence of mutually exclusive events \( E_1, E_2, \ldots \) (i.e. events where \( E_i E_j = \emptyset \) when \( i \neq j \))

\[
P \left( \bigcup_{i=1}^{\infty} E_i \right) = \sum_{i=1}^{\infty} P(E_i)
\]

If \( P(F) > 0 \), then
\[
P(E \mid F) = \frac{P(EF)}{P(F)}
\]

\[
P(A \cup B) = P(A) + P(B) - P(AB)
\]

\[
P(E) = P(EF) + P(EF^c)
\]

\[
= P(E \mid F)P(F) + P(E \mid F^c)P(F^c)
\]

\[
= P(E \mid F)P(F) + P(E \mid F^c)[1 - P(F)]
\]

\[
P(E_1 E_2 E_3 \ldots E_n) = P(E_1)P(E_2 \mid E_1)P(E_3 \mid E_2 E_1) \ldots P(E_n \mid E_1 \ldots E_{n-1})
\]

\[
P(E) = P(EF) + P(EF^c)
\]

\[
= P(E \mid F)P(F) + P(E \mid F^c)P(F^c)
\]

\[
= P(E \mid F)P(F) + P(E \mid F^c)[1 - P(F)]
\]

\[
P(E_1 E_2 E_3 \ldots E_n) = P(E_1)P(E_2 \mid E_1)P(E_3 \mid E_2 E_1) \ldots P(E_n \mid E_1 \ldots E_{n-1})
\]
**Definition**

*independent events*

**Probability**

Definition

*probability mass function of a discrete random variable*

Theorem

**Definition**

*cumulative distribution function $F$*

**Probability**

Theorem

*properties of the cumulative distribution function*

**Definition**

*expected value (discrete case)*

**Proposition**

*expected value of a function of $X$ (discrete case)*

**Corollary**

**Definition/Theorem**

*linearity of expectation*

**Definition**

*variance*

**Definition**

*probability mass function of a Bernoulli random variable*

**Probability**

*probability mass function of a binomial random variable*
For a discrete random variable $X$, we define the probability mass function $p(a)$ of $X$ by

$$p(a) = P\{X = a\}$$

Probability mass functions are often written as a table.

Two events $E$ and $F$ are said to be independent iff

$$P(EF) = P(E)P(F)$$

Otherwise they are said to be dependent.

The cumulative distribution function satisfies the following properties:

1. $F$ is a nondecreasing function
2. $\lim_{a \to \infty} F(a) = 1$
3. $\lim_{a \to -\infty} F(a) = 0$

The cumulative distribution function $(F)$ is defined to be

$$F(a) = \sum_{all \ x \leq a} p(x)$$

The cumulative distribution function $F(a)$ denotes the probability that the random variable $X$ has a value less than or equal to $a$.

If $X$ is a discrete random variable that takes on the values denoted by $x_i$ ($i = 1 \ldots n$) with respective probabilities $p(x_i)$, then for any real–valued function $f$

$$E[f(X)] = \sum_{i=1}^{n} f(x_i)p(x)$$

If $X$ is a random variable with mean $\mu$, then we define the variance of $X$ to be

$$\text{var}(X) = E[(X - \mu)^2] = E[X^2] - (E[X])^2 = E[X^2] - \mu^2$$

If $\alpha$ and $\beta$ are constants, then

$$E[\alpha X + \beta] = \alpha E[X] + \beta$$

The first line is the actual definition, but the second and third equations are often more useful and can be shown to be equivalent by some algebraic manipulation.

Suppose $n$ independent Bernoulli trials are performed. If the probability of success is $p$ and the probability of failure is $1 - p$, then $X$ is said to be a binomial random variable with parameters $(n, p)$.

The probability mass function is given by:

$$p(i) = \binom{n}{i} p^i (1 - p)^{n-i}$$

where $i = 0, 1, \ldots, n$

If an experiment can be classified as either success or failure, and if we denote success by $X = 1$ and failure by $X = 0$ then, $X$ is a Bernoulli random variable with probability mass function:

$$p(0) = P\{X = 0\} = 1 - p$$

$$p(1) = P\{X = 1\} = p$$

where $p$ is the probability of success and $0 \leq p \leq 1$. 
A random variable \( X \) that takes on one of the values \( 0, 1, \ldots \), is said to be a Poisson random variable with parameter \( \lambda \) if for some \( \lambda > 0 \)

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

where \( i = 0, 1, 2, \ldots \)

Suppose independent Bernoulli trials, are repeated until success occurs. If we let \( X \) equal the number of trials required to achieve success, then \( X \) is a geometric random variable with probability mass function:

\[
p(n) = P\{X = n\} = (1 - p)^{n-1}p
\]

where \( n = 1, 2, \ldots \)

Suppose that independent Bernoulli trials (with probability of success \( p \)) are performed until \( r \) successes occur. If we let \( X \) equal the number of trials required, then \( X \) is a negative binomial random variable with probability mass function:

\[
p(n) = P\{X = n\} = \binom{n - 1}{r - 1} p^r (1 - p)^{n-r}
\]

where \( n = r, r + 1, \ldots \)

We define \( X \) to be a continuous random variable if there exists a function \( f \), such that for any set \( B \) of real numbers

\[
P\{X \in B\} = \int_B f(x) \, dx
\]

The function \( f \) is called the probability density function of the random variable \( X \).

If \( X \) is a binomial random variable with parameters \( n \) and \( p \), then

\[
E[X] = np
\]

\[
\text{var}(X) = np(1 - p)
\]

If \( X \) is a Poisson random variable with parameter \( \lambda \), then

\[
E[X] = \lambda
\]

\[
\text{var}(X) = \lambda
\]

If \( X \) is a negative binomial random variable with parameters \( (p, r) \), then

\[
E[X] = \frac{r}{p}
\]

\[
\text{var}(X) = \frac{r(1 - p)}{p^2}
\]

If \( X \) is a geometric random variable with parameter \( p \), then

\[
E[X] = \frac{1}{p}
\]

\[
\text{var}(X) = \frac{1 - p}{p^2}
\]

If \( X \) is a uniform random variable on the interval \( (\alpha, \beta) \), then its probability density function is given by

\[
f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{if } \alpha < x < \beta \\ 0 & \text{otherwise} \end{cases}
\]

If \( X \) is a uniform random variable with parameters \((\alpha, \beta)\), then

\[
E[X] = \frac{\alpha + \beta}{2}
\]

\[
\text{var}(X) = \frac{(\beta - \alpha)^2}{12}
\]