Joint, Conditional, and Total Probabilities; Independence

Joint Probability

\[ \Pr(A, B) \]

Independent events A and B

\[ \Pr(A, B) = \Pr(B, A) = \Pr(A) \cdot \Pr(B) \]

Intersecting events (not independent)

\[ \Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) \]

Intersecting events (independent)

\[ \Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A) \cdot \Pr(B) \]

Intersecting events – conditional probability

\[ \Pr(A \cap B) = \Pr(A \mid B) \cdot \Pr(B), \text{ for } \Pr(B) > 0 \]
\[ \Pr(A \cap B) = \Pr(B \mid A) \cdot \Pr(A), \text{ for } \Pr(A) > 0 \]
\[ \Pr(A \cap B) = \Pr(A \mid B) \cdot \Pr(B) = \Pr(B \mid A) \cdot \Pr(A), \text{ for } \Pr(A) > 0, \Pr(B) > 0 \]

Conditional Probability

\[ \Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A, B)}{\Pr(B)}, \text{ for } \Pr(B) > 0 \]

Conditional Probability – A and B independent

\[ \Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A) \cdot \Pr(B)}{\Pr(B)} = \Pr(A), \text{ for } \Pr(B) > 0 \]

Conditional Probability – \( A \subseteq B \)

\[ \Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A)}{\Pr(B)}, \text{ for } \Pr(B) > 0 \]

Conditional Probability – \( B \subseteq A \)

\[ \Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(B)}{\Pr(B)} = 1, \text{ for } \Pr(B) > 0 \]

Conditional Probability – A and B mutually exclusive

\[ \Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{0}{\Pr(B)} = 0, \text{ for } \Pr(B) > 0 \]
**Total Probability**

\[ S = A_1 \cup A_2 \cup A_3 \cdots \cup A_n \]

\[ A_i \cap A_j = \emptyset, \text{ for } i \neq j \]

then

\[ B = B \cap S = B \cap (A_1 \cup A_2 \cup A_3 \cdots \cup A_n) = (B \cap A_1) \cup (B \cap A_2) \cup (B \cap A_3) \cdots \cup (B \cap A_n) \]

\[ \Pr(B) = \Pr(B \cap A_1) + \Pr(B \cap A_2) + \Pr(B \cap A_3) \cdots + \Pr(B \cap A_n) \]

\[ \Pr(B) = \Pr(B \mid A_1) \cdot \Pr(A_1) + \Pr(B \mid A_2) \cdot \Pr(A_2) + \cdots + \Pr(B \mid A_n) \cdot \Pr(A_n) \]

**Bayes Theorem (a-priori used to derive a-posteriori probability)**

\[ \Pr(A_i \cap B) = \Pr(B \mid A_i) \cdot \Pr(A_i) = \Pr(A_i \mid B) \cdot \Pr(B) \]

\[ \Pr(A_i \mid B) = \frac{\Pr(B \mid A_i) \cdot \Pr(A_i)}{\Pr(B)}, \text{ for } \Pr(B) > 0 \]

\[ \Pr(A_i \mid B) = \frac{\Pr(B \mid A_i) \cdot \Pr(A_i)}{\Pr(B \mid A_1) \cdot \Pr(A_1) + \Pr(B \mid A_2) \cdot \Pr(A_2) + \cdots + \Pr(B \mid A_n) \cdot \Pr(A_n)} \]

**Bernoulli Trials**

The probability that an event occurs \( k \) times in \( n \) independent trials of an experiment can be defined as

\[ \Pr(\text{A occurring } k \text{ times in } n \text{ trials}) = p_n(k) = b(k; n, p) = \binom{n}{k} \cdot p^k \cdot q^{n-k} \]

Note that

Binomial coefficient have the mathematical notations \( _n C_k \) or \( \binom{n}{k} \) or \( C(n, k) \).

Note: a useful relationship

\[ \binom{n}{k} = \binom{n}{n-k} = \frac{n!}{k!(n-k)!} \]

And another math sum

\[ \sum_{k=0}^{n} \binom{n}{k} = 2^n \]
Poisson probability mass function (pmf) and approximation from future chapters ...

Conditions: \( n \gg 1, \quad p \ll 1, \quad k \ll n \), but \( n \cdot p = \mu \) is a constant term, then

\[
b(k; n, p) = p_n(k) = \binom{n}{k} p^k q^{n-k} \approx \frac{1}{k!} \cdot \mu^k \cdot \left(1 - \frac{\mu}{n}\right)^{n-k}
\]

Then, if we consider an infinite number of trials \( \ldots \) that is \( n \to \infty \)

\[
b(k; n, p) \approx \frac{1}{k!} \cdot \mu^k \cdot \left(1 - \frac{\mu}{n}\right)^{n-k} \approx \frac{\mu^k}{k!} \cdot \exp(-\mu)
\]

Normal Approximation to the Binomial Law (DeMoivre-Laplace Theorem)

This is an approximation of the binomial distribution when the number of trials \( n \) is large and other assumptions are met.

Assumptions: \( n \cdot p \cdot q \gg 1 \) and \( |k - n \cdot p| \leq \sqrt{n \cdot p \cdot q} \)

\[
p_n(k) = \binom{n}{k} \cdot p^k \cdot q^{n-k} \approx \frac{1}{\sqrt{2 \pi \cdot n \cdot p \cdot q}} \cdot \exp\left(-\frac{(k - n \cdot p)^2}{2 \cdot n \cdot p \cdot q}\right)
\]

[Aside: we will be discussing the law of large numbers and that sums of larger numbers of events appear as a Gaussian distribution. This is the first example \( \ldots \) and you haven't been told what a Gaussian distribution is yet.]

Binomial Law

“Summation of Bernoulli trials” \( \Rightarrow k \) or fewer successes in \( n \) trials

\[
B(k; n, p) = \sum_{i=0}^{k} b(i; n, p) = \sum_{i=0}^{k} \binom{n}{i} \cdot p^i \cdot q^{n-i}
\]