

# Stats Reference

*Kelly Sovacool*

*2018-11-19*



# Contents

<b>Preface</b>	<b>5</b>
<b>1 Calculus Review</b>	<b>7</b>
1.1 Logarithms . . . . .	7
1.2 Derivative & Integration rules . . . . .	7
1.3 Trigonometry . . . . .	8
<b>2 Probability</b>	<b>9</b>
2.1 Axioms . . . . .	9
2.2 Union Rule . . . . .	9
2.3 Inclusion-Exclusion Principle . . . . .	9
2.4 De Morgan's Laws . . . . .	9
2.5 Conditional Probability . . . . .	9
2.6 Bayes' Theorem . . . . .	10
2.7 Independence . . . . .	10
2.8 Counting Examples . . . . .	10
<b>3 Distributions of Random Variables</b>	<b>11</b>
3.1 Discrete . . . . .	11
3.2 Continuous . . . . .	12
3.3 Properties . . . . .	14
3.4 Distributions of Functions . . . . .	15
<b>4 Joint Distributions</b>	<b>17</b>
4.1 Marginals . . . . .	17
4.2 Independence . . . . .	17
4.3 Sums of Independent Random Variables . . . . .	17
4.4 Conditional Joint Distributions . . . . .	18
4.5 Order Statistics . . . . .	19
4.6 Transformations of Joint Distributions . . . . .	19
<b>5 Expectation</b>	<b>21</b>
5.1 Expectation $\mathbb{E}[x]$ . . . . .	21
5.2 Variance $\text{Var}[x]$ . . . . .	21
5.3 Covariance $\text{Cov}[x, y]$ . . . . .	21
5.4 Correlation Coefficient $\rho$ . . . . .	21



# Preface

I started compiling this reference for concepts & equations in statistics while taking an introductory probability theory course at [U-M](#).

You can view the [bookdown site](#) or download the [PDF file](#).

The source code is hosted on [GitHub](#).



# Chapter 1

## Calculus Review

### 1.1 Logarithms

$$\log_b a = x \leftrightarrow b^x = a$$

$$e^{c \ln x} = x^c$$

### 1.2 Derivative & Integration rules

Derivative	Integral
$\frac{d}{dx} x^n = nx^{n-1}$	$\int x^n dx = \frac{x^{n+1}}{n+1} + c$
$\frac{d}{dx} n^x = n^x \ln n$	$\int n^x dx = \frac{n^x}{\ln n} + c$
$\frac{d}{dx} \ln x = \frac{1}{x}$	$\int \frac{1}{ax+b} dx = \frac{1}{a} \ln  ax+b  + c$
$\frac{d}{dx} e^x = e^x$	$\int e^x dx = e^x + c$
$\frac{d}{dx} \sin x = \cos x$	$\int \sin x dx = -\cos x + c$
$\frac{d}{dx} \cos x = -\sin x$	$\int \cos x dx = \sin x + c$
$\frac{d}{dx} \tan x = \sec^2 x$	$\int \tan x = \ln  \sec x  + c$

$$\int f(x) \pm g(x) dx = \int f(x) dx \pm \int g(x) dx$$

$$\int x f(x) = x F(x) + f(x)$$

#### 1.2.1 Quotient Rule

$$\frac{d}{dx} \left( \frac{f(x)}{g(x)} \right) = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}$$

#### 1.2.2 Integration by substitution

$$u = g(x)$$

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(u)du$$

### 1.2.3 Integration by parts

Assign  $u$  and  $dv$ , differentiate  $u$  to find  $du$ , integrate  $dv$  to find  $v$ , then solve:

$$\int_a^b u dv = [uv]_a^b - \int_a^b v du$$

## 1.3 Trigonometry

### 1.3.1 SOH CAH TOA

### 1.3.2 Basic Identities

### 1.3.3 Pythagorean Identities

$$\sin^2 x + \cos^2 x = 1$$

$$\tan^2 x + 1 = \sec^2 x$$

$$1 + \cot^2 x = \csc^2 x$$



## Chapter 2

# Probability

For the following section,  $A$  and  $B$  represent events in the sample space  $S$ .

### 2.1 Axioms

1.  $\mathbb{P}(A) \geq 0 \quad \forall A \subset S$
2.  $\mathbb{P}(S) = 1$
3. If  $A \cap B = \emptyset$ , then  $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

### 2.2 Union Rule

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

### 2.3 Inclusion-Exclusion Principle

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)$$

### 2.4 De Morgan's Laws

$$(A \cup B)^c = A^c \cap B^c$$

$$(A \cap B)^c = A^c \cup B^c$$

### 2.5 Conditional Probability

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

$$\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c)$$

## 2.6 Bayes' Theorem

$$\mathbb{P}(B_j|A) = \frac{\mathbb{P}(A|B_j)\mathbb{P}(B_j)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_j)\mathbb{P}(B_j)}{\sum_{i=1}^k \mathbb{P}(A|B_i)\mathbb{P}(B_i)}$$

## 2.7 Independence

If events  $A$  and  $B$  are independent:  $\mathbb{P}(A|B) = \mathbb{P}(A)$

## 2.8 Counting Examples

- There are  $n!$  ways to arrange  $n$  distinct elements in an ordered list.
- There are  $6^n$  outcomes for  $n$  tosses of a 6-sided die.

## Chapter 3

# Distributions of Random Variables

### 3.1 Discrete

$$\text{CDF: } F(k) = \sum_{k=0}^i p(k)$$

#### 3.1.1 Bernoulli

$$X \sim \text{Bern}(p)$$

$$\mathbb{E}[X] = p$$

$$\text{Var}[X] = p(1 - p)$$

$$p(x) = \begin{cases} p & x = 1 \\ 1 - p & x = 0 \\ 0 & \text{else} \end{cases} \quad (3.1)$$

#### 3.1.2 Binomial

$$X \sim \text{Binom}(n, p)$$

$$\mathbb{E}[X] = np$$

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

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

#### 3.1.3 Poisson

$$X \sim \text{Poisson}(\lambda \sim np)$$

$$\mathbb{E}[X] = \lambda$$

$$\text{Var}[X] = \lambda$$

$$p(k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

- Approximation to binomial when  $n \rightarrow \infty$  and  $p \rightarrow 0$ .
- E.g. number of misprints per page of a book.

### 3.1.4 Geometric

$$X \sim \text{Geom}(p)$$

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

$$\text{Var}[X] = \frac{1-p}{p^2}$$

$$p(k) = (1-p)^{k-1} \tag{3.2}$$

$$F(k) = 1 - (1-p)^k$$

- Experiment with infinite trials; stop at first success.
- Memoryless.
- E.g. flip a coin until heads comes up.

### 3.1.5 Hypergeometric

pmf

exp

var

## 3.2 Continuous

### 3.2.1 Uniform

$$X \sim \text{Unif}(a, b)$$

$$\mathbb{E}[X] = \frac{b+a}{2}$$

$$\text{Var}[X] = \frac{(b-a)^2}{12}$$

$$f(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{else} \end{cases} \tag{3.3}$$

$$F(x) = \begin{cases} \frac{x-a}{b-a} & x \in [a, b] \\ 0 & \text{else} \end{cases}$$

- Simplest continuous distribution.
- All outcomes equally likely.
- E.g. uniformly pick random point on disk of radius  $r$ .  $x$  is distance to center (not Uniform).  $f(x) = \frac{2x}{r^2}$  when  $0 \leq x \leq r$ .

### 3.2.2 General Normal

$$X \sim N(\mu, \sigma)$$

$$\mathbb{E}[X] = \mu$$

$$\text{Var}[X] = \sigma^2$$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} \quad (3.4)$$

$$F(x) = \Phi\left(Z = \frac{x-\mu}{\sigma}\right)$$

- To find CDF, convert to standard normal, then use Z table.
- E.g. biological variables.
- E.g. exam scores.

### 3.2.3 Standard Normal

$$X \sim N(0, 1)$$

$$\mathbb{E}[X] = 0$$

$$\text{Var}[X] = 1$$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-(x^2)/2} \quad (3.5)$$

$$F(x) = \Phi(x)$$

- To find CDF, use Z table.
- Special case of the normal with no parameters.

#### 3.2.3.1 Normal Approximation to the Binomial Distribution

When  $X \sim \text{Binom}(n, p)$ ,  $n \rightarrow \infty$ , &  $p \rightarrow \frac{1}{2}$ :

$$\mathbb{E}[X] = np = \mu, \sigma = \sqrt{np(1-p)}, z = \frac{x-np}{\sqrt{np(1-p)}}$$

$$F_z(a) \rightarrow \Phi(a)$$

$$\mathbb{P}(a \leq z \leq b) \approx \Phi(b) - \Phi(a)$$

via De Moivre-Laplace Central Limit Theorem

### 3.2.4 Exponential

$$X \sim \text{Exp}(\lambda)$$

$$\mathbb{E}[X] = \frac{1}{\lambda}$$

$$\text{Var}[X] = \frac{1}{\lambda^2}$$

$$\begin{aligned}
 f(x) &= \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{else} \end{cases} \\
 F(x) &= \begin{cases} 1 - e^{-\lambda x} & x > 0 \\ 0 & \text{else} \end{cases}
 \end{aligned} \tag{3.6}$$

- Memoryless.
- $\lambda$  = rate.
- Continuous version of  $\text{Geom}(p)$ .

### 3.2.4.1 Hazard & Survival

Survival:  $S_T(t) = \mathbb{P}(T > t) = 1 - \mathbb{P}(T \leq t) = 1 - F_T(t) = e^{-\int_{u=0}^t \lambda(u) du}$

Density:  $f_T(t) = F'_T(t) = -S'_T(t)$

Hazard:  $\lambda(t) = \frac{f_T(t)}{S_T(t)} = \frac{-S'_T(t)}{S_T(t)} = -\frac{d}{dt} \log S_T(t)$

### 3.2.5 Gamma

$X \sim \Gamma[\alpha, \lambda]$

$\mathbb{E}[X] = \frac{\alpha}{\lambda}$

$\text{Var}[X] = \frac{\alpha}{\lambda^2}$

### 3.2.6 Chi Square

$X \sim \chi^2[n]$

$\mathbb{E}[X] = n$

$\text{Var}[X] = 2n$

- Special case of  $\Gamma$  where  $\alpha = \frac{n}{2}$  and  $\lambda = \frac{1}{2}$ .

## 3.3 Properties

### 3.3.1 Density Functions

PMF:  $p(k)$     PDF:  $f(x)$

- Derivative of the distribution function.
- Nonnegative everywhere.
- Integral over its domain is 1:  $\int_a^b f(x) = 1$

### 3.3.2 Distribution Functions

CDF:  $F(x)$

- $\lim_{x \rightarrow -\infty} F(x) = 0$
- $\lim_{x \rightarrow +\infty} F(x) = 1$
- Nondecreasing.

### 3.3.3 Parameters

Law of total expectation:  $\mathbb{E}[X] = \sum_j \mathbb{E}(E|F_j)\mathbb{P}(F_j)$

Discrete:  $\mathbb{E}[X] = \mu = \sum_{i=1}^k x_i p_i$

Continuous:  $\mathbb{E}[X] = \mu = \int_{-\infty}^{\infty} x f(x) dx$  (3.7)

$$\text{Var}[X] = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \sigma^2$$

$$\sigma = \sqrt{\text{Var}[X]}$$

$$Z = \frac{x - \mu}{\sigma}, \quad Z \sim N(0, 1)$$

## 3.4 Distributions of Functions

$X$  is a random variable.  $Y = g(x)$  is a function of  $X$ .

### 3.4.1 Transformation Method

If  $Y = g(x)$  is monotonic:

$$f_Y(y) = \frac{1}{|g'(x)|} f_X(x)$$

Derive  $g'(x)$  from  $g(x)$ . Integrate  $f_Y$  to find  $F_Y$ .

Note: monotonic = invertible = one-to-one.

### 3.4.2 CDF Method

Must use when  $Y = g(x)$  is not monotonic:

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(g(x) \leq y) \rightarrow \text{solve for } x \text{ and frame in terms of } F_X(y).$$

Differentiate  $F_Y$  to find  $f_Y$ .





## Chapter 4

# Joint Distributions

$$\mathbb{P}(x \in A, y \in B) = \int_A \int_B f(x, y) dy dx$$

$$\mathbb{E}[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dx dy$$

### 4.1 Marginals

$$f_X = \int f(x, y) dy$$

$$f_Y = \int f(x, y) dx$$

$$\mathbb{P}(x \in A) = \mathbb{P}(x \in A, y \in (-\infty, \infty)) = \int_A \int_{-\infty}^{\infty} f(x, y) dy dx$$

$$\mathbb{P}(y \in B) = \mathbb{P}(x \in (-\infty, \infty), y \in B) = \int_{-\infty}^{\infty} \int_B f(x, y) dy dx$$

### 4.2 Independence

$$f(x, y) = f_X(x) f_Y(y) \quad \forall x, y$$

$$F(x, y) = F_X(x) F_Y(y) \quad \forall x, y$$

#### 4.2.1 Minimum & Maximum

$$\text{Max: } F_{\text{Max}}(t) = \mathbb{P}(\text{Max} \leq t) = \mathbb{P}(x \leq t, y \leq t) \rightarrow \text{use independence} \rightarrow = F_X(t) F_Y(t)$$

$$\text{Min: } 1 - F_{\text{Max}}$$

### 4.3 Sums of Independent Random Variables

#### 4.3.1 Distributions

$$\text{Convolution (CDF): } F_{X+Y}(a) = \mathbb{P}(X + Y \leq a) = \int_{-\infty}^{\infty} F_X(a - y) f_Y(y) dy$$

$$\text{Density (PDF): } f_{X+Y} = \int_{-\infty}^{\infty} f_X(a - y) f_Y(y) dy$$

### 4.3.2 Uniform

### 4.3.3 Normal

The sum of  $n$  normal RVs  $\sum_i^n X_i$  is normally distributed with parameters:

$$\begin{aligned}\mu &= \sum_i^n \mu_i \\ \sigma^2 &= \sum_i^n \sigma_i^2 \\ \sigma &= \sqrt{\sum_i^n \sigma_i^2} \neq \sum_i^n \sqrt{\sigma_i^2}\end{aligned}$$

### 4.3.4 Gamma

### 4.3.5 Poisson

$$X_1 \sim \text{Poisson}(\lambda_1)$$

$$X_2 \sim \text{Poisson}(\lambda_2)$$

$$Y = X_1 + X_2$$

$$Y \sim \text{Poisson}(\lambda = \lambda_1 + \lambda_2)$$

$$\mathbb{P}(X_1 + X_2 = n) = \frac{1}{n!} e^{-(\lambda_1 + \lambda_2)} (\lambda_1 + \lambda_2)^n$$

### 4.3.6 Binomial

$$X_1 \sim \text{Binom}(n, p)$$

$$X_2 \sim \text{Binom}(m, p)$$

$$Y = X_1 + X_2$$

$$Y \sim \text{Binom}(n + m, p)$$

$$\mathbb{P}(X_1 + X_2 = k) = \binom{n+m}{k} = \sum_{i=0}^n \binom{n}{i} \binom{m}{k-i}$$

### 4.3.7 Geometric

## 4.4 Conditional Joint Distributions

### 4.4.1 Discrete

$$P_{X|Y} = \frac{P(x,y)}{P_Y(y)} = \mathbb{P}(X = x|Y = y)$$

$$\mathbb{E}[X|Y = y] = \sum_x x P_{X|Y}(x|y)$$

### 4.4.2 Continuous

$$f_{X|Y} = \frac{f(x,y)}{f_Y(y)}$$

$$\mathbb{E}[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx$$

$$F_{X|Y}(a, y) = \mathbb{P}(X \leq a|Y = y) = \int_{-\infty}^a f_{X|Y}(x|y) dx$$

### 4.4.3 Bayes' Theorem (Continuous)

$$f_{X|Y} = \frac{f_{Y|X}(y|x)f_x(x)}{f_Y(y)} = \frac{f_{Y|X}(y|x)f_x(x)}{\int f_{Y|X}(y|x)f_x(x) dx}$$

## 4.5 Order Statistics

## 4.6 Transformations of Joint Distributions

### 4.6.1 The Jacobian

$$(u, v) = G(x, y)$$

$$\text{Jac}(x, y) = \det \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} = \frac{\partial u}{\partial x} \frac{\partial v}{\partial y} - \frac{\partial u}{\partial y} \frac{\partial v}{\partial x}$$

$$f_{u,v}(u, v) = \frac{1}{|\text{Jac}(x, y)|} f_{x,y}(x, y)$$



## Chapter 5

# Expectation

### 5.1 Expectation $\mathbb{E}[x]$

Law of total expectation:  $\mathbb{E}[X] = \sum_j \mathbb{E}(E|F_j)\mathbb{P}(F_j)$

Discrete:  $\mathbb{E}[X] = \mu = \sum_{i=1}^k x_i p_i$

Continuous:  $\mathbb{E}[X] = \mu = \int_{-\infty}^{\infty} x f(x) dx$

Conditional expectation:

$$\mathbb{E}[X] = \mathbb{E}(\mathbb{E}[X|Y])$$

$$\mathbb{E}[XY] = \mathbb{E}(\mathbb{E}[XY|Y]) = \mathbb{E}(Y\mathbb{E}[X|Y])$$

### 5.2 Variance $\text{Var}[x]$

where  $a$  is a constant:

$$\text{Var}(ax) = a^2 \text{var}(x)$$

$$\text{Var}(a) = 0$$

### 5.3 Covariance $\text{Cov}[x, y]$

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

### 5.4 Correlation Coefficient $\rho$

$$\rho(x, y) = \frac{\text{Cov}(x, y)}{\sqrt{\text{Var}[X]}\sqrt{\text{Var}[Y]}}$$