Appendix B

Notation

This appendix contains a summary of the notation used in this book.

Math notation

Expression	Read as	Used to denote
$=$	is equal to	expressions that have the same value
$\stackrel{\text{def}}{=}$	is defined as	a new variable definition
$a + b$	a plus b	the combined lengths of a and b
$a - b$	a minus b	the difference between a and b
$a \cdot b \stackrel{\text{def}}{=} ab$	a times b	the area of a rectangle
$a^2 = aa$	a squared	the area of a square of side length a
$a^3 = aaa$	a cubed	the volume of a cube of side length a
a^n	a exponent n	a multiplied by itself n times
$\sqrt{a} = a^{\frac{1}{2}}$	square root of a	the side length of a square of area a
$a/b = \frac{a}{b}$	a divided by b	a parts of a whole split into b parts
$a^{-1} = \frac{1}{a}$	one over a	division by a
$\%$	percent	proportions of a total; $a\% \stackrel{\text{def}}{=} \frac{a}{100}$
$f(x)$	f of x	the function f applied to input x
f^{-1}	f inverse	the inverse function of f
e^x	e to the x	the exponential function base e
$\ln(x)$	natural log of x	the logarithm base e
$\lceil x \rceil$	ceiling of x	round up x to nearest integer
$\lfloor x \rfloor$	floor of x	round down x to nearest integer

Set notation

Symbol	Read as	Denotes
$\{\ldots\}$	the set ...	define a sets
$\mid$	such that	describe or restrict the elements of a set
$\mathbb{N}$	the naturals	the set $\mathbb{N} \stackrel{\text{def}}{=} \{0,1,2,\ldots\}$. Note $\mathbb{N}_+ \stackrel{\text{def}}{=} \mathbb{N}\backslash\{0\}$
$\mathbb{N}_+$	the positive naturals	the natural without zero $\mathbb{N}_+ \stackrel{\text{def}}{=} \{1,2,3,\ldots\}$
$\mathbb{Z}$	the integers	the set $\mathbb{Z} \stackrel{\text{def}}{=} \{\ldots,-2,-1,0,1,2,3,\ldots\}$
$\mathbb{Q}$	the rationals	the set of fractions of integers
$\mathbb{R}$	the reals	the set of real numbers
$\subset$	subset	one set strictly contained in another
$\subseteq$	subset or equal	containment or equality
$\cup$	union	the combined elements from two sets
$\cap$	intersection	the elements two sets have in common
$S\backslash T$	S set minus T	the elements of S that are not in T
$a \in S$	a in S	a is an element of set S
$a \notin S$	a not in S	a is not an element of set S
$\forall x$	for all x	a statement that holds for all x
$\exists x$	there exists x	an existence statement
$\nexists x$	there doesn't exist x	a non-existence statement
$\neg A$	logical NOT	is true when A is false
$A \vee B$	logical OR	is true when either A or B are true
$A \wedge B$	logical AND	is true when both A and B are true

Data notation

Symbol	Read as	Denotes
Mean	mean	average value
Med	median	middle value of the dataset
		Ideas for median notation: $x_{0.5}$ $x_{\frac{1}{2}}$ $\bar{\bar{x}}$
Mode	mode	the most frequently observed value
Var	variance	average squared deviation from mean
Std	standard deviation	the square root of the variance
$\mathbf{Q}_1$	first quartile	one quarter of values smaller than $\mathbf{Q}_1$
$\mathbf{Q}_2 = \mathbf{Med}$	second quartile	middle value of the dataset
$\mathbf{Q}_3$	third quartile	one quarter of values are larger than $\mathbf{Q}_3$
IQR	interquartile range	span of the middle fifty percent of the data
Min	minimum	the smallest value in a dataset
Max	maximum	the largest value in a dataset
Range	range	difference between the **Max** and **Min**
$\mathbf{Cov}(x, y)$	covariance	measures the joint variability of x and y
$\mathbf{Corr}(x, y)$	correlation	measures the linear relatedness of x and y
$\mathbf{Freq}(v)$	frequency	number of v in a dataset
$\mathbf{RelFreq}(v)$	relative frequency	number of v in a dataset divided by n

Calculus notation

Expression	Denotes
$f(x)$	a function of the form $f : \mathbb{R} \to \mathbb{R}$
$f'(x)$	derivative of $f(x)$
$\frac{d}{dx}$	derivative operator
$\int_a^b f(x)\,dx$	integral of $f(x)$ between $x = a$ and $x = b$
$F(x)$	the integral function of $f(x)$
a_k	sequence $a_k : \mathbb{N} \to \mathbb{R}$, also denoted $(a_0, a_1, a_2, a_3, \ldots)$
$\sum_{k=r}^{s} a_k$	summation of the terms $a_r + a_{r+1} + \cdots + a_{s-1} + a_s$

Probability notation

Expression	Denotes
$\{\text{descr}\}$	an outcome described by the condition "descr"
$\Pr(\{\text{descr}\})$	the probability of the outcome $\{\text{descr}\}$
$\Pr(\{\text{descr}\}\vert\{\text{cond}\})$	conditional probability of outcome $\{\text{descr}\}$ given $\{\text{cond}\}$
$\mathbb{1}_{\{\text{descr}\}}$	indicator function for the outcome $\{\text{descr}\}$
X	a random variable (abbreviated as r.v.)
$\mathcal{X}$	the set of possible outcomes for the r.v. X
x	a particular outcome of the r.v. X
$f_X(x) \stackrel{\text{def}}{=} \Pr(\{X = x\})$	*probability mass function* (pmf) of a discrete r.v. X, or *probability density function* (pdf) of a continuous r.v. X.
$X \sim f_X$	X is distributed according to f_X
$F_X(x) \stackrel{\text{def}}{=} \Pr(\{X \leqslant x\})$	cumulative distribution function (CDF) of the r.v. X
$\mathbb{E}_X[\cdot]$	expectation operator with respect to r.v. X
$\mu \stackrel{\text{def}}{=} \mathbb{E}_X[X]$	the *mean* of X
$\sigma^2 \stackrel{\text{def}}{=} \mathbb{E}_X[(X-\mu)^2]$	the *variance* of X; also denoted $\mathbf{var}(X)$
$\sigma = \sqrt{\sigma^2}$	the *standard deviation* of X
$\mathcal{N}(\mu, \sigma)$	the *normal* distribution with mean μ and standard deviation σ. The probability density function for the r.v. $X \sim \mathcal{N}(\mu, \sigma)$ is $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$
$Z \sim \mathcal{N}(0, 1)$	the *standard normal distribution*
$\Gamma(z)$	the Gamma function $\Gamma(z) \stackrel{\text{def}}{=} \int_{y=0}^{y=\infty} y^{z-1}e^{-y}dy$
$\mathbf{X} \stackrel{\text{def}}{=} (X_1, X_2, \ldots, X_n)$	random sample of size n. Each $X_i \sim f_X$
$\mathbb{E}_{\mathbf{X}}[\cdot]$	expectation with respect to random sample $\mathbf{X}$
$X \perp Y$	X and Y are independent
$X \not\perp Y$	X and Y are not independent
$\mathbf{x} \stackrel{\text{def}}{=} (x_1, x_2, \ldots, x_n)$	sample of n observations
$f_{\mathbf{x}}$	empirical pmf of sample $\mathbf{x} = (x_1, x_2, \ldots, x_n)$
$F_{\mathbf{x}}$	empirical CDF of sample $\mathbf{x} = (x_1, x_2, \ldots, x_n)$
$F_{\mathbf{x}}^{-1}$	inverse of the empirical CDF of the sample $\mathbf{x}$

Probability distributions reference

Name	Math notation	Code	Parameters	Notes
Discrete uniform	$\mathcal{U}_d$	`randint(alpha, beta+1)`	$\alpha, \beta \in \mathbb{Z}$	
Bernoulli	Bernoulli	`bernoulli(p)`	$p \in [0,1]$	
Binomial	Binom	`binom(n,p)`	$n \in \mathbb{N}_+$, $p \in [0,1]$	
Poisson	Pois	`poisson(lam)`	$\lambda \in \mathbb{R}_+$	
Geometric	Geom	`geom(p)`	$p \in (0,1)$	
Negative binomial	NBinom	`nbinom(r,p)`	$r \in \mathbb{N}_+$, $p \in (0,1)$	(optiona
Hypergeometric	Hypergeom	`hypergeom(a+b,a,n)`	$a, b, n \in \mathbb{N}_+$	(optiona
Multinomial	Multinomial	`multinomial`	$n \in \mathbb{N}_+$, $p_i \in [0,1]$	(optiona
Uniform	$\mathcal{U}$	`uniform`	$\alpha, \beta \in \mathbb{R}$	
Exponential	Expon	`expon`	$\lambda \in \mathbb{R}_+$	
Normal	$\mathcal{N}$	`norm`	$\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}_+$	...
Standard normal	$Z \sim \mathcal{N}(0,1)$	`norm(0,1)`		...
Gamma	Gamma	`gamma`		
Beta	Beta	`beta`		
Student's T	$\mathcal{T}$	`t`		
Chi square	χ^2	`chi2`		
Snedecor's F	$\mathcal{F}$	`f`		
Cauchy	Cauchy	`cauchy`		(optiona
Laplace	Laplace	`laplace`		(optiona
Pareto	Pareto	`pareto`		(optiona
Wald	Wald	`wald`		(optiona
Weibull	Weibull	`weibull_min`		(optiona
Zipf	Zipf	`zipf`		(optiona

Each of these models provides you with the same set of methods for computing probabilities, confidence intervals, and generating random values: `.pmf(x)` or `.pdf(x)`, `.cdf(b)`, `.expect(f)`, `.mean()`, `.std()`, `.rvs(n)`, etc.

TODO: show math params for all distributions, and equivalent params for computer models `scipy.stats`

Statistics notation

Expression	Denotes
f_X	*probability mass function* of a discrete r.v. X, or *probability density function* of a continuous r.v. X.
n	sample size
$\mathbf{x} \stackrel{\text{def}}{=} (x_1, x_2, \ldots, x_n)$	a particular sample of size n from the r.v. X
$\mathbf{X} \stackrel{\text{def}}{=} (X_1, X_2, \ldots, X_n)$	random sample of size n. Each $X_i \sim f_X$
$\hat{\theta} = g(\mathbf{x})$	particular value of the estimator g computed from the sample $\mathbf{x} = (x_1, x_2, \ldots, x_n)$
$\widehat{\Theta} = g(\mathbf{X})$	sampling distribution of the estimator g computed from the random sample $\mathbf{X} = (X_1, X_2, \ldots, X_n)$
$\overline{\mathbf{x}}$	sample mean
$s_{\mathbf{x}}^2$	sample variance
$s_{\mathbf{x}}$	sample standard deviation
$\mathbf{se}_g$	standard error of the estimator g
$\widehat{\mathbf{se}}_g$	estimated standard error of the estimator g
θ	a parameter of the probability distribution
$\hat{\theta}$	an estimate of the parameter θ
μ	population mean
σ^2	population variance
σ	population standard deviation
ν	degrees of freedom parameter
CV_g	cutoff value for a decision based on estimator g

Linear models notation

Expression	Denotes
$\tilde{\mathbf{x}}$	new, previously unseen input
$\tilde{y}$	predicted output value for the input $\tilde{\mathbf{x}}$