Notes on Notation

We generally follow notational standards common in applied statistics. But to a student, notation can often prove a barrier. This notation "glossary" is meant to ease the transition to reading notation-heavy material and provide a place to look up unfamiliar symbols. The underlying assumption is that students have already been introduced to basic probability, calculus, and linear algebra concepts.

Random variables and sets are denoted using script capitals. Thus, for example, $X = \{\ldots, -2, 0, 2, \ldots\}$ denotes the set of even integers. $Y \sim f_N(0, 1)$ states that Y is random variable that is distributed according to a Gaussian normal distribution with mean of 0 and variance of 1. We will denote the set of admissible values for X (its support) as $\mathcal{X}$.

Both upper- and lowercase letters can represent functions. When both upper- and lowercase versions of the same letter are used, the uppercase function typically represents the integral of the lowercase function, e.g., $G(x) = \int_{-\infty}^{x} g(u)du$.

To conserve notation we will use $f_s(\cdot; \theta)$ to represent the probability distribution and mass functions commonly used in building Generalized Linear Models. θ denotes generic parameters, possibly vector-valued. The subscript will denote the specific distribution:

- f_B is the Bernoulli distribution
- f_b is the binomial distribution
- f_β is the Beta distribution
- f_c is the categorical distribution
- f_e is the exponential distribution
- f_{EV_1} is the type-I extreme value distribution
- f_Γ is the Gamma distribution
- f_{GEV} is the generalized extreme value distribution
- f_L is the logistic distribution

https://doi.org/10.1017/9781316888544.003 Published online by Cambridge University Press

- f_{lL} is the log-logistic distribution
- f_m is the multinomial distribution
- $f_{\mathcal{N}}$ is the Gaussian (Normal) distribution
- f_{Nb} is the negative binomial
- f_P is the Poisson distribution
- F_W is the Weibull distribution

To conform with conventional terminology and notation in $\mathcal{R}$, we refer to one-dimensional vectors as *scalars*. Scalars and observed realizations of random variables are denoted using lowercase math script. $\Pr(Y_i \leq y_i)$ denotes the probability that some random variable, Y_i, takes a value no greater than some realized level, y_i.

Matrices are denoted using bolded capital letters; $\mathbf{X}_{n\times k}$ is the matrix with n rows and k columns. The symbol $\intercal$ denotes matrix or vector transposition, as in $\mathbf{X}^\intercal$. Vectors are represented with bolded lowercase letters, e.g., $\mathbf{x}_i$. In our notation we implicitly treat all vectors as *column* vectors unless otherwise stated. For example, $\mathbf{x}_i$ is a column vector even though it may represent a row in the $\mathbf{X}_{n\times k}$ matrix. "Barred" items denote the sample mean e.g., $\bar{\mathbf{y}}$.

Lowercase Greek letters are typically reserved for parameters of models and statistical distributions. These parameters could be either scalars or vectors. Vectors will be expressed in bold font. Where more specificity is needed we will subscript.

"Hatted" objects denote fitted or estimated quantities; when used in the context of an MLE then hatted objects are the MLE. For example, β might be a regression parameter and $\hat{\beta}$ is the estimated value of that parameter.

Common functions, operators, and objects:

- $\propto$ means "is proportional to"
- $\dot{\sim}$ means "approximately distributed as"
- $\xrightarrow{d}$ means "convergence in distribution."
- $\xrightarrow{p}$ means "convergence in probability," what some texts denote plim.
- $\mathbb{1}(\cdot)$ is the indicator function that returns a 1 if true and a 0 otherwise.
- $\mathrm{cov}(\cdot,\cdot)$ is the covariance function
- $\det(\cdot)$ is the determinant of a square matrix
- $\mathrm{E}[\cdot]$ is the expectation operator
- $\exp(\cdot)$ is the exponential function
- $\Gamma(\cdot)$ is the Gamma function
- $\mathbf{I}_n$ is the $n \times n$ identity matrix
- $\mathcal{I}(\cdot)$ is the expected Fisher information
- $I(\cdot)$ is the observed Fisher information
- *iid* means "independently and identically distributed."
- $\Lambda(\cdot)$ is the logistic cumulative distribution function
- log is the logarithm. If no base is given, then it denotes the natural logarithm (base e)

https://doi.org/10.1017/9781316888544.003 Published online by Cambridge University Press

- ∇ is the gradient vector of some function.
- $\Phi(\cdot)$ is the standard Normal cumulative distribution function
- $\phi(\cdot)$ is the standard Normal density function
- $\Pr(\cdot)$ denotes probability
- $\text{var}(\cdot)$ is the variance function

https://doi.org/10.1017/9781316888544.003 Published online by Cambridge University Press