Glossary

$100(1-\alpha)\%$ confidence interval a continuous interval (L,U) with lower bound L and upper bound U satisfying $Pr(L \leq \theta \leq U) = 1-\alpha$. 107

k-component spliced distribution a distribution whose PDF is given by piecing together other PDFs over disjoint intervals of the sample space. 43

n-fold convolution (denoted $F_X^{*n}(x)$) the distribution defined by

$$F_X^{*0}(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

and

$$F_X^{*k}(x) = \int_{-\infty}^{\infty} F_X^{*(k-1)}(x-y) dF_X(y)$$

for $k = 1,2,\ldots$. 85

p-value the probability of observing data at least as extreme as your test statistic under the null hypothesis. 110

0-1 loss a loss function defined by $l(\theta,\hat{\theta}) = \begin{cases} 1 & if\ \theta \neq \hat{\theta} \\ 0 & if\ \theta = \hat{\theta} \end{cases}$. 200

100p-th percentile any value π_p such that the CDF $F(\pi_p^-) \leq p \leq F(\pi_p)$. 11

$(\mathbf{a},\mathbf{b},\mathbf{0})$ class a class of discrete distributions such that the PMF p_k satisfies $\frac{p_k}{p_{k-1}} = a + \frac{b}{k}$, for constants a and b and for $k \in \{1,2,3,\ldots\}$. 52

$(\mathbf{a},\mathbf{b},\mathbf{1})$ class a class of discrete distributions such that the PMF p_k satisfies $\frac{p_k}{p_{k-1}} = a + \frac{b}{k}$, for constants a and b and for $k \in \{2,3,\ldots\}$. 54

absolute margin of error half the width of the confidence interval. 247

Anderson-Darling test a hypothesis test that is used to test whether the data are consistent with a particular model CDF. 224

asymptotically unbiased property of an estimator $\hat{\theta}_n$ (which depends on the sample size n) whose *bias* decreases to 0 as n increases to ∞. 104

bias (denoted $Bias_{\hat{\theta}}(\theta)$) a measure of how far away the estimator $\hat{\theta}_n$ tends to be from the true value θ, calculated as $E(\hat{\theta}|\theta) - \theta$. 103

censored from above see *right censored*. 127

censored from below see *left censored*. 127

Central Limit Theorem (CLT) Let $\{X_1,\ldots X_k\}$ be a sequence of random variables, and let S_k denote their sum, i.e. $S_k = X_1 + X_2 + \ldots + X_k$. Under appropriate conditions,

$$\lim_{k\to\infty} \frac{S_k - E(S_k)}{\sqrt{Var(S_k)}} \rightsquigarrow N(0,1)$$

where $\rightsquigarrow$ means convergence in distribution. 18

coefficient of variation the standard deviation divided by the mean. 13

coherent describes a risk measure that satisfies the properties of subadditivity, monotonicity, positive homogeneity, and translation invariance. 28

coinsurance (denoted α) the proportion of a loss covered by the insurance company, after applying inflation and the deductible. 64

complement of credibility $1-z$ where z is *credibility*. 266

complete expectation of life see *excess loss function*. 23

compound Poisson distribution the distribution of $S = X_1 + \cdots + X_N$, where frequency N is Poisson is distributed. 83

conditional tail expectation see *tail value-at-risk*. 28

conditional value-at-risk see *tail value-at-risk*. 28

consistent a property of an estimator $\hat{\theta}_n$ such that $\lim_{n\to\infty} Pr(|\hat{\theta}_n - \theta| > \delta) = 0$ for all $\delta > 0$ and for all θ. 104

continuous distribution a distribution which assigns positive probability to a set of continuous values. 8

Cramer-Rao inequality under typical regularity conditions, no unbiased estimator has a variance smaller than $1/I(\theta)$. 171

credibility (denoted z) a value between 0 and 1 that is the weight we give to our estimated value derived from observed data. 266

credibility-weighted estimate an estimate E derived from a weighted average of an estimate derived from observed data and another estimate, where the weight on the observation-derived estimate is *credibility* z. 266

cumulative distribution function (CDF) $F(x) := Pr(X \leq x)$. 8

cumulative hazard function (denoted $H(x)$ for a random variable X with hazard function $h(x)$, $H(x) := \int_{-\infty}^{x} h(y)dy$. 116

difference plot only available for individual data, a plot of the difference between the empirical distribution and a target distribution $D(x) = F_n(x) - F^*(x)$ on the y-axis against observed values x on the x-axis. 214

discrete distribution a distribution which assigns positive probability to a set of discrete values. 8

distribution describes the relationship between a set of values and their associated probabilities. 8

empirical distribution see *empirical model*. 10

empirical model a discrete distribution derived from an observed dataset of n data points $\{x_1, \ldots, x_n\}$, where each observed value is assigned a probability of $1/n$. 10

event a set of outcomes that occur with some probability. 8

excess loss function the expected value of a *excess loss variable*. 23

excess loss variable see *per-payment random variable*. 23

expectation of a random variable the weighted average of all possible values that a random variable can take on. 11

expected cost of stop-loss insurance (denoted $E[(S-d)_+]$) the expected value of S with a deductible of d. 86

expected shortfall see *tail value-at-risk*. 28

expected value see *expectation of a random variable*. 11

expected value of the process variance (EPV) the expected value of the variances in each subgroup $E_\theta\left(Var(X|\theta)\right)$. 286

franchise deductible (denoted d) a threshold d for which there is no payout if $X \leq d$ and a payout of X if $X > d$. 60

full credibility criterion a threshold n_0 on sample size n such that $n > n_0$ implies that credibility $z = 1$. 266

hazard function denoted $h(x)$), given by $\frac{f(x)}{S(x)}$. 22

heavy-tailed describes a distribution with high probabilities concentrated at large values of the random variable. 26

improper prior distribution a prior distribution that does not integrate to 1. 198

indicator random variable (denoted $\mathbf{1}\{X \leq x\}$) defined as

$$\mathbf{1}\{X \leq x\} = \begin{cases} 1 & X \leq x \\ 0 & X > x \end{cases}$$

where X is a random variable. 248

information (denoted $I(\theta)$) calculated as $I(\theta) = E_\theta\left(\left[\frac{\partial}{\partial\theta}\log L(\theta|x)\right]^2\right)$ in the one-dimensional case, with a similar formula in the multi-dimensional case. 167

inverse transform sampling a way to sample values from a parametric distribution by making use of simulations from a Uniform$(0,1)$ distribution. 244

inversion sampling see *inverse transform sampling*. 244

joint distribution (for data x and parameter θ) $f_{X,\Theta}(x,\theta) = f_{X|\Theta}(x|\theta)\cdot\pi(\theta)$. 198

Kaplan-Meier product-limit (KM) estimator an estimator of the survival function $S(t)$ for modified data. 130

kernel (denoted $k_y(x)$) the underlying distribution used in constructing a kernel density estimator. 152

kernel density estimator a nonparametric estimator of the CDF and PDF of a data set, created by centering a kernel distribution around each data point. 152

Kolmogorov-Smirnov Test sometimes abbreviated as K-S test, a hypothesis test that is used to test whether the data are consistent with a particular model CDF. 223

k-point mixture a random variable Y constructed from random variables $X_1,\ldots,X_k$ such that $F_Y(y) = a_1F_{X_1}(y) + \cdots + a_kF_{X_k}(y)$ where $\sum_{i=1}^k a_i = 1$. 33

k-th central moment (of a random variable X) given by $E[(X-\mu)^k]$, sometimes denoted μ_k. 13

k-th raw moment (of a random variable X) given by $E[X^k]$, sometimes denoted μ'_k. 12

kurtosis $\gamma_2 := \frac{E[(X-\mu)^4]}{\sigma^4}$. 13

left censored value of the random variable is recorded as is when it exceeds the censoring point d and is recorded as d when it is at or below d. 127

left truncated value of the random variable is recorded as is when it exceeds the truncation point d and not recorded when it is at or below d. 127

left-censored-and-shifted variable see *per-loss random variable*. 24

left-skewed distribution has negative *skewness*, oftentimes evidenced in a elongated left tail. 15

left-truncated-and-shifted variable see *per-payment random variable*. 23

light-tailed describes a distribution with low probabilities concentrated at large values of the random variable. 26

likelihood function (denoted $L(\theta|x)$) a function of the underlying parameters θ that that is equal to the probability of the observed data x given they are distributed with parameters x. 162

likelihood ratio test a hypothesis test that is used to test whether the data are more consistent with a more restrictive class of models compared to a superset class of models. 229

limited expected value (denoted $E(X \wedge u)$) the expected value of a *limited loss random variable*. 24

limited loss random variable (denoted $X \wedge u$) a random variable modeling payout which has a maximum payment of u. 24

linear exponential family a family of distributions whose PDFs take the form of $f(x;\theta) = \frac{p(x)e^{r(\theta)x}}{q(\theta)}$, where $p(x)$ does not depend on θ and $r(\theta)$ and $q(\theta)$ do not depend on x. 43

loss elimination ratio (LER) percent decrease in the expected payment with a policy modification, calculated as $\frac{E(X)-E(Y^L)}{E(X)}$. 62

loss function a user-defined function used to measure the error of the estimate $\hat{\theta}$ given that the true parameter is θ. 200

marginal density a PDF for a subset of all of the available random variables, often obtained by integrating (or summing) the joint density over the random variables that are not included. 40

marginal distribution (for data x and parameter θ) $f_X(x) = \int f_{X|\Theta}\pi(\theta)d\theta$. 198

maximum covered loss (denoted u) the largest amount of your losses that will be covered. 64

maximum likelihood estimator (MLE) (denoted $\hat{\theta}_{MLE}$) the estimator of θ which maximizes the likelihood $L(\theta|x)$. 162

mean see *expectation of a random variable*. 11

mean excess loss function see *excess loss function*. 23

mean residual life function see *excess loss function*. 23

mean squared error (MSE) a measure of the bias and variance of an estimator $\hat{\theta}$ calculated as $MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2|\theta]$. 106

median the 50th percentile $\pi_{0.5}$. 11

method of moments (MOM) (denoted $\hat{\theta}_{MOM}$) an estimator of the parameters of a distribution obtained by matching theoretical moments with empirical moments. 159

model distribution (denoted $f_{X|\Theta}(x|\theta)$, also known as the likelihood) the probability distribution for the data x under θ, a specific $\theta \in \Theta$. 198

moment generating function (MGF) (denoted $M_X(t)$) equal to $E(e^{tx})$. 15

$M_X(t)$ see *moment generating function*. 15

Nelson-Åalen (NA) estimator an estimator of the cumulative hazard function, calculated by

$$\hat{H}(x) = \begin{cases} 0 & 0 \le x < y_1 \\ \sum_{i=1}^{j-1} \frac{s_i}{r_i} & y_{j-1} \le x < y_j, \quad 2 \le j \le k \\ \sum_{i=1}^{k} \frac{s_i}{r_i} & x \ge y_k \end{cases}$$

for unique, ordered observation times $y_1, \ldots, y_k$. 117

net stop-loss premium see *expected cost of stop-loss insurance*. 86

observed information an estimator of information $I(\theta)$ constructed by plugging in actual observed data rather than taking the expected value. 173

ogive (denoted $F_n(x)$) an estimated CDF for grouped data. 119

ordinary deductible (denoted d) a threshold d for which there is no payout if $X \leq d$ and a payout of $X - d$ if $X > d$. 59

p-p plot only available for individual data, a plot of the points of the empirical and the model CDF, where the empirical CDF is evaluated with a denominator of $n+1$. 218

partial credibility the value of credibility $0 \leq z < 1$, which occurs when sample size $n < n_0$, where n_0 is the *full credibility criterion*. 266

per-loss random variable (denoted Y^L) a random variable modeling payout, where the random variable equals 0 if no payout occurred (for example, if losses were below the deductible). 24

per-payment random variable (denoted Y^P) a random variable taking on strictly positive values used to model payout conditional on the existence of the payout being larger than 0. 23

percentile matching an estimator of the parameters of a distribution obtained by matching theoretical percentiles with empirical percentiles. 160

policy limit (denoted u) the maximum payout, after the application of a deductible. 63

posterior distribution (for data x and parameter θ) $\pi_{\Theta|X}(\theta|x) = \frac{f_{X|\Theta}(x|\theta)\cdot\pi(\theta)}{\int f_{X|\Theta}(x|\theta)\pi(\theta)d\theta}$. 198

power one minus the *type II error*. 111

predictive distribution (denoted $f_{Y|X}(y|x)$) distribution for predicting the next observation given n observed values and the posterior distribution, calculated as $\int f_{Y|\Theta}(y|\theta)\pi_{\Theta|X}(\theta|x)d\theta$. 202

prior distribution a PDF for the prior random variable, usually a random variable that is a parameter in the distribution of another random variable. 40, 198

probability density function (PDF) (denoted $f(x)$ or $p(x)$) defined for a continuous random variable X, it expresses the relative probability of $X = x$. 8

probability generating function (PGF) (denoted by $P_X(z)$) equal to $E(z^X)$. 15

probability mass function (PMF) (denoted $f(x)$ or $p(x)$) defined for a discrete random variable X, it expresses the probability of $X = x$. $p(x) := Pr(X = x)$. 8

process variance (denoted σ^2_{PP}) variance of *pure premiums*. 273

pure premium (PP) total severity divided by the number of exposures. 266

$P_X(z)$ see *probability generating function*. 15

random variable a variable which can take on multiple values. 8

relative margin of error a value c such that $a = (1-c)\theta$ and $b = (1+c)\theta$ where θ is the parameter for which the confidence interval (a,b) is constructed. 247

right censored value of the random variable is recorded as is when it is below the censoring point u and is recorded as u when it is at or above u. 127

right truncated value of the random variable is recorded as is when it is below the truncation point u and not recorded when it is at or above u. 127

right-censored variable see *limited loss random variable*. 24

right-skewed distribution has positive *skewness*, oftentimes evidenced in a elongated right tail. 15

sample space a set of values that a random variable can take. 8

scale distribution a distribution for which multiplying all values in the sample space by a constant $c > 0$ still yields the same distributional form (with possibly different parameters). 22

Schwarz-Bayesian Criterion a score-based model selection approach in which log-likelihoods are adjusted by being penalized for model complexity and then compared. 232

significance level see *type I error*. 109

skewness $\gamma_1 := \frac{E[(X-\mu)^3]}{\sigma^3}$. 13

standard deviation (denoted σ) the square root of the *variance*. 13

standard for full credibility see *full credibility criterion*. 266

standard normal distribution a normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$. 14

survival function (denoted $S(x)$), given by $Pr(X > x) = 1 - F(x)$. 22

tail value-at-risk (TVaR) expected loss given that the loss exceeds the $100p$-th *percentile* of the distribution of X. 30

truncated from above see *right truncated*. 127

truncated from below see *left truncated*. 127

type I error (denoted α) the probability of rejecting H_0 when it is in fact true (a false positive). 109

type II error (denoted β) probability of failing to reject H_0 when H_a is true (a false negative). 109

unbiased property of an estimator $\hat{\theta}_n$ whose *bias* is equal to zero. 103

uniformly minimum variance unbiased estimator (UMVUE) an unbiased estimator $\hat{\theta}$ such that for any θ, there is no other unbiased estimator that has smaller variance. 107

value-at-risk (VaR) (denoted $VaR_p(x)$ or $\pi_p(x)$) equal to the $100p$-th *percentile* of the distribution of X. 29

variable-component mixture a *k-point mixture* where k is not predetermined. 33

variance (denoted σ^2, $Var(X)$) the second central moment of X, given by $E[(X-\mu)^2)]$. 13

variance of the hypothetical means (VHM) the variance of all the expected values in each subgroup $Var_\theta\left(E\left(X|\theta\right)\right)$. 286

$(X-x)_+$ called the positive part of $X-x$, computed as follows:

$$(X-x)_+ := \begin{cases} X-x & X > x \\ 0 & X \leq x \end{cases}$$

Namely, a function of X which returns 0 if $X \leq x$ and $X-x$ otherwise. 86

Y^L see *per-loss random variable*. 24

Y^P see *per-payment random variable*. 23

${}_{y-x}p_x$ probability of living past time y given survival up to time x. 5, 132

${}_{y-x}q_x$ probability of dying before time y given survival up to time x. 5, 132

zero-modified for the $(\mathbf{a},\mathbf{b},\mathbf{1})$ class, a modification of the original PMF by setting $p_0 = Pr(X = 0) > 0$. 54

zero-truncated for the $(\mathbf{a},\mathbf{b},\mathbf{1})$ class, a modification of the original PMF by setting $p_0 = Pr(X = 0) = 0$. 54