Serik Sagitov, Chalmers and GU, January 25, 2018

An overview of the Probability Theory

1 Probability rules

Sample space:

Ω = the set of all possible outcomes in a random experiment.

Random events:

$A, B \subset \Omega$, $A \cap B = \{A \text{ and } B\}$, $A \cup B = \{A \text{ or } B \text{ or both}\}$

Division rule:

$\mathrm{P}(A) = \frac{\text{no. favorable outcomes}}{\text{total no. equally likely outcomes}}$

Addition rule:

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

Complementary event probability:

$\mathrm{P}(\bar{A}) = 1 - \mathrm{P}(A)$, $\bar{A} = \{A \text{ has not occurred}\}$

Conditional probability of A given B has occurred:

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

Multiplication rule:

Independent events: $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$

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

Law of Total Probability:

if $\{B_1, \ldots, B_k\}$ is a partition of Ω, then $\mathrm{P}(A) = \sum_{i=1}^{k} \mathrm{P}(B_i)\mathrm{P}(A|B_i)$

Bayes' Probability Law

prior probabilities $\mathrm{P}(B_i)$ and posterior probabilities $\mathrm{P}(B_i|A) = \frac{\mathrm{P}(A|B_i)\mathrm{P}(B_i)}{\mathrm{P}(A)}$

2 Random variables

Discrete random variable X:

probability mass function $f(x) = \mathrm{P}(X = x)$

Continuous random variable X:

probability density function $f(x) \approx \frac{\mathrm{P}(x<X<x+\Delta)}{\Delta}$

Cumulative distribution function (cdf)

$$F(x) = \mathrm{P}(X \le x) = \sum_{y \le x} f(y) \text{ or } = \int_{y \le x} f(y)dy.$$

Mean (average or expected) value of X:

$$\mu = \mathrm{E}(X) = \sum_x x f(x) \quad \text{or} \quad \mu = \int x f(x)dx.$$

Properties of the expectation operator

$$\mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y), \quad \mathrm{E}(c \cdot X) = c \cdot \mathrm{E}(X).$$

Variance and standard deviation

$$\sigma_X^2 = \mathrm{Var}(X) = \mathrm{E}((X-\mu)^2) = \mathrm{E}(X^2) - \mu^2, \quad \sigma_X = \sqrt{\mathrm{Var}(X)},$$
$$\mathrm{E}(X^2) = \sum x^2 f(x) \text{ or } \mathrm{E}(X^2) = \int x^2 f(x)dx,$$
$$\mathrm{Var}(c \cdot X) = c^2 \cdot \mathrm{Var}(X), \quad \sigma_{cX} = c \cdot \sigma_X.$$

3 Special distributions

Discrete uniform distribution $X \sim \mathrm{dU}(N)$:

$$f(k) = \frac{1}{N}, \quad 1 \le k \le N, \quad \mathrm{E}(X) = \frac{N+1}{2}, \quad \mathrm{Var}(X) = \frac{N^2-1}{12}.$$

Continuous uniform distribution $X \sim \mathrm{U}(a,b)$:

$$f(x) = \frac{1}{b-a}, \quad a < x < b, \quad \mathrm{E}(X) = \frac{a+b}{2}, \quad \mathrm{Var}(X) = \frac{(b-a)^2}{12}.$$

Binomial distribution $X \sim \mathrm{Bin}(n,p)$:

$$f(k) = \binom{n}{k} p^k q^{n-k}, \quad 0 \le k \le n, \quad \mathrm{E}(X) = np, \quad \mathrm{Var}(X) = npq.$$

Hypergeometric distribution $X \sim \mathrm{Hg}(N,n,p)$:

$$f(k) = \frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}}, \quad \max(0, n-Nq) \le k \le \min(n, Np),$$
$$\mathrm{E}(X) = np, \quad \mathrm{Var}(X) = npq\Big(1 - \frac{n-1}{N-1}\Big),$$

where the factor $1 - \frac{n-1}{N-1}$ is called the finite population correction.

Geometric distribution $X \sim \mathrm{Geom}(p)$:

$$f(k) = pq^{k-1}, \quad k \ge 1, \quad \mathrm{E}(X) = \frac{1}{p}, \quad \mathrm{Var}(X) = \frac{q}{p^2}$$

Exponential distribution $X \sim \text{Exp}(\lambda)$:

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0, \quad \text{E}(X) = \sigma_X = \frac{1}{\lambda}$$

Poisson distribution $X \sim \text{Pois}(\lambda)$:

$$f(k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k \geq 0, \quad \text{E}(X) = \text{Var}(X) = \lambda$$

Poisson approximation

$$\text{Bin}(n, p) \approx \text{Pois}(np) \quad \text{if } n \geq 100 \text{ and } p \leq 0.01 \text{ and } np \leq 20$$

Standard normal distribution $Z \sim \text{N}(0, 1)$:

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad \Phi(z) = \int_{-\infty}^{z} \phi(x)dx,$$
$$\text{E}(Z) = 0, \quad \text{Var}(Z) = 1.$$

Normal distribution $X \sim \text{N}(\mu, \sigma^2)$:

$$\frac{X - \mu}{\sigma} \sim \text{N}(0, 1), \quad f(x) = \frac{1}{\sigma}\phi(\frac{x - \mu}{\sigma}), \quad \text{E}(X) = \mu, \quad \text{Var}(X) = \sigma^2$$

Central Limit Theorem (CLT): if $X_1, \ldots, X_n$ are IID, $\text{E}(X_i) = \mu$, $\text{Var}(X_i) = \sigma^2$, then for all z,

$$\text{P}\left(\frac{X_1 + \ldots + X_n - n\mu}{\sigma\sqrt{n}} \leq z\right) \to \Phi(z), \quad n \to \infty.$$

Normal approximations:

$\text{Bin}(n, p) \approx \text{N}(np, npq)$, if both $np \geq 5$ and $nq \geq 5$,

$\text{Pois}(\lambda) \approx \text{N}(\lambda, \lambda)$, for $\lambda \geq 5$,

$\text{Hg}(N, n, p) \approx \text{N}(np, npq\frac{N-n}{N-1})$, if both $np \geq 5$ and $nq \geq 5$.

4 Joint distributions

Joint probability mass (density) function of X and Y:

$$f_{X,Y}(x, y)$$

Marginal distribution of X:

$$f_X(x) = \sum_y f_{X,Y}(x, y) \quad \text{or} \quad \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy$$

Independent random variables X and Y:

$$f_{X,Y}(x, y) = f_X(x)f_Y(y)$$

Conditional distribution of $(Y|X=x)$:

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

Conditional expectation $\mathrm{E}(Y|X)$ and $\mathrm{Var}(Y|X)$

$$\mathrm{E}(\mathrm{E}(Y|X)) = \mathrm{E}(Y)$$
$$\mathrm{Var}(Y) = \mathrm{Var}(\mathrm{E}(Y|X)) + \mathrm{E}(\mathrm{Var}(Y|X))$$

Addition rule for variance

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y)$$

Covariance, a measure of association between X and Y

$$\mathrm{Cov}(X,Y) = \mathrm{E}(X-\mu_X)(Y-\mu_Y) = \mathrm{E}(XY) - \mu_X\mu_Y$$

Correlation coefficient

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y}, \quad -1 < \rho < 1$$

Independent X and Y are uncorrelated: $\rho = 0$. If X and Y are uncorrelated, then

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

Multinomial distribution $(X_1,\ldots,X_r) \sim \mathrm{Mn}(n; p_1,\ldots,p_r)$

$$\mathrm{P}(X_1 = k_1,\ldots,X_r = k_r) = \binom{n}{k_1,\ldots,k_r} p_1^{k_1}\ldots p_r^{k_r}, \quad \mathrm{Cov}(X_i,X_j) = -np_ip_j$$

Bivariate normal distribution $(X,Y) \sim \mathrm{N}(\mu_X,\mu_Y,\sigma_X^2,\sigma_Y^2,\rho)$:

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left\{\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right\}\right)$$

Marginal distributions :

$$X \sim \mathrm{N}(\mu_X,\sigma_X^2), \quad Y \sim \mathrm{N}(\mu_Y,\sigma_Y^2)$$

Conditional distribution of $(Y|X=x)$ is normal with mean and variance

$$\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X), \qquad \sigma_y^2(1-\rho^2)$$