Formula Page

Algebra of set theory

Union: $C = A \cup B$; **Intersection**: $C = A \cap B$; **Complement**: $C = A^c$
Commutative Laws: $A \cup B = B \cup A$, $A \cap B = B \cap A$
Associative Laws: $(A \cup B) \cup C = A \cup (B \cup C)$, $(A \cap B) \cap C = A \cap (B \cap C)$
Distributive Laws: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$
De Morgan's Law: $(A \cap B)^c = A^c \cup B^c$, $(A \cup B)^c = A^c \cap B^c$

Computing probabilities

k samples from n objects

	ordered	unordered
replacement	n^k	$\frac{(n+k-1)!}{(n-1)!\,k!}$
no replacement	$\frac{n!}{(n-k)!}$	$\frac{n!}{k!\,(n-k)!}$

Conditional probabilities、Bayes' Theorem、Law of total probability

Addition Law: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Multiplication Law: $P(A \cap B) = P(A|B)P(B)$
Bayes' theorem: $P(B|A) = \frac{P(A|B)P(B)}{P(A)}$
Law of total probability: $P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_n)P(B_n)$
Type of probabilities:

Events	B_1	…	B_n
Prior probabilities	$P(B_1)$	…	$P(B_n)$
Conditional probabilities	$P(A|B_1)$	…	$P(A|B_n)$
Joint probabilities	$P(A|B_1)P(B_1)$、$P(A \cap B_1)$	…	$P(A|B_n)P(B_n)$、$P(A \cap B_n)$
Posterior probabilities	$P(B_1|A)$	…	$P(B_n|A)$

Independence

$$\text{Pairwise independence} \not\Rightarrow\Leftarrow \text{Mutual independence}$$

Discrete rv

$E(X) = \sum_i x_i \cdot p_X(x_i) = \int_{-\infty}^{+\infty} x f_X(x) dx$
$Y = f(X)$, $E(Y) = \sum_i f(x_i) \cdot p_X(x_i)$
$E(cX + d) = cE(X) + d$
$Var(X) = E[(X-\mu)^2] = \sum_i (x_i - \mu)^2 p_X(x_i) = \int_{-\infty}^{+\infty} (x-\mu)^2 f_X(x) dx = E(X^2) - [E(X)]^2$
$Var(cX + d) = c^2 Var(X)$
Bernoulli distribution: X takes on only two values e.g. 0 and 1, with probabilities $1 - p$ and p.
$p(x) = \begin{cases} p^x(1-p)^{1-x}, & \text{if } x = 0 \text{ or } x = 1 \\ 0, & \text{otherwise} \end{cases}$, Mean: p, Variance: $p(1-p)$
Binomial distribution: X is the sum of n variables that follow Bernoulli distribution.
$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$, Mean: np, Variance: $np(1-p)$
Limit: Poisson Distribution, $n \to \infty, p \to 0, np \nrightarrow \infty, X \sim \text{Poisson}(np)$
Geometric distribution (total number of trials): X is the total trials until the first success.
$P(X = k) = (1-p)^{k-1} p$, Mean: $1/p$, Variance: $(1-p)/p^2$
Geometric distribution (number of failures): X is the total fail trials until the first success.

$P(X=k)=(1-p)^k p$, Mean: $(1-p)/p$, Variance: $(1-p)/p^2$

Poisson Distribution: X is the number of times an event occurs in a given interval of time.

$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$, Mean:$\lambda$, Variance: λ

Limit: Normal Distribution, $\lambda \to \infty, X \sim N(\lambda, \lambda)$

Continuous rv

pth quantile $\mathbf{x_p}$: $F(x_p)=P(X \le x_p)=p$

Uniform Distribution:

pdf: $f_X(x)=\begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0, & x < a \text{ or } x > b \end{cases}$, cdf: $F_X(x)=\begin{cases} 0, x \le a \\ \frac{x-a}{b-a}, a < x < b \\ 1, x \ge b \end{cases}$, Mean: $\frac{a+b}{2}$, Variance: $\frac{(a-b)^2}{12}$

Normal distribution: $X \sim N(\mu, \sigma^2)$, pdf: $f_X(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}$, Mean: μ, Variance: σ^2

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

$$(X,Y) \sim N_2(\boldsymbol{\mu}, \Sigma) \Rightarrow\not\Leftarrow X \sim N(\mu_X, \sigma_X^2) \text{ and } Y \sim N(\mu_Y, \sigma_Y^2)$$

Exponential distribution: the time until next event. λ: count per time unit

pdf: $f_X(x)=\lambda\exp(-\lambda x), x \ge 0, \lambda > 0$, cdf: $F_X(x)=1-\exp(-\lambda x), x \ge 0$

Mean: $1/\lambda$, Variance: $1/\lambda^2$

χ^2 distribution: pdf: $f_X(x)=\frac{1}{\Gamma(p/2)2^{p/2}}x^{\frac{p}{2}-1}e^{-x/2}, x > 0$; $Z \sim N(0,1)$, $Z^2 \sim \chi_1^2$

Joint Distribution

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy$$

$P(x_1 < X \le x_2, y_1 < Y \le y_2)=F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)$

Covariance

$Cov(X,Y)=E(XY)-E(X)E(Y)$

$Cov(aW+bX, cY+dZ)=acCov(W,Y)+adCov(W,Z)+bcCov(X,Y)+bdCov(Y,Z)$

$Var(aX+bY)=a^2Var(X)+b^2Var(Y)+2abCov(X,Y)$

$$\text{independent} \Rightarrow\not\Leftarrow \rho(X,Y)=0$$

Limiting Theorems

Chebyshev's Inequality:

Let X be a random variable with known $E(X)$ and $Var(X)=\sigma^2$ but unknown distribution

$$P(|X-E(X)| \ge k\sigma) \le \frac{1}{k^2}$$

Law of Large numbers:

The sample mean will be close to μ if the sample size is sufficiently large.

Central Limit Theorem:

Let $X_1, X_2, \dots, X_n$ be a sequence of independent random variables having a common distribution with mean μ and variance σ^2. We can compute probabilities for the sample mean using the cdf of the Normal distribution:

$$\overline{X} \sim N(\mu, \sigma^2/n)$$