FORMULAE

Version July, 2017

Probability and Statistics

401-2604-00, Spring 2017

1. Standard discrete distributions

(1) Bernoulli distribution with success parameter $p \in (0,1)$. $X \in \{0,1\}$ and

$$P(X=1) = p, \quad EX = p, \quad \mathrm{Var}(X) = p(1-p).$$

(2) Binomial distribution with n trials and success parameter $p \in (0,1)$. $X \in \{0,1,\ldots,n\}$

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0,1,\ldots n,$$

$$EX = np, \quad \mathrm{Var}(X) = np(1-p).$$

(3) Poisson distribution with parameter $\lambda > 0$. $X \in \{0,1,\ldots\}$

$$P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0,1,\ldots,$$

$$EX = \lambda, \quad \mathrm{Var}(X) = \lambda.$$

2. Standard continuous distributions

(4) Gaussian distribution with mean μ and variance σ^2. $X \in \mathbb{R}$,

$$f_X(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \quad x \in \mathbb{R}.$$

Denoted by $X \sim \mathcal{N}(\mu, \sigma^2)$.

$$EX = \mu, \ \mathrm{var}(X) = \sigma^2.$$

$$X \sim \mathcal{N}(\mu,\sigma^2) \quad \Leftrightarrow \quad Z := \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1).$$

$\mathcal{N}(0,1)$ is called the standard normal (or Gaussian).

(5) The standard normal distribution function.

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-z^2/2}\, dz, \quad x \in \mathbb{R}.$$

Let Φ^{-1} be its inverse function. Then,

$$\Phi^{-1}(0.9) = 1.28, \quad \Phi^{-1}(0.95) = 1.64, \quad \Phi^{-1}(0.975) = 1.96.$$

(6) Exponential distribution with parameter $\lambda > 0$. $X \in \mathbb{R}_+ := [0,\infty)$,

$$f_X(x) = \frac{1}{\lambda} e^{-x/\lambda}, \quad x \geq 0.$$

$$EX = \lambda, \quad \mathrm{Var}(X) = \lambda^2.$$

Note: in many textbooks λ is replaced by $1/\lambda$.

(7) Gamma distribution with parameters α, λ. $X \in \mathbb{R}_+ := [0, \infty)$,

$$f_X(x) = \frac{1}{\lambda^\alpha \Gamma(\alpha)} \, x^{\alpha - 1} \, e^{-x/\lambda}, \quad x \geq 0.$$

Here $\Gamma(\alpha)$ is the Gamma function and for integer values $\Gamma(m) = (m-1)!$.

$$EX = \alpha\lambda, \quad \mathrm{Var}(X) = \alpha\lambda^2.$$

Note: in many textbooks λ is replaced by $1/\lambda$.

(8) Beta distribution with parameters r, s. $X \in [0, 1]$,

$$f_X(x) = \frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} \, x^{r-1} \, (1-x)^{s-1}, \quad x \in [0, 1].$$

$$EX = \frac{r}{r+s}, \quad \mathrm{Var}(X) = \frac{rs}{(r+s)^2 \, (1+r+s)}.$$

(9) Chi-Square (χ^2) distribution.
The χ^2 distribution with m degrees of freedom is the Gamma distribution with parameters $\alpha = m/2, \ \lambda = 2$. Denoted by $\chi^2(m)$. In particular,

$$X \sim \mathcal{N}(0, 1) \quad \Rightarrow \quad X^2 \sim \chi^2(1),$$

$$X_j \sim \mathcal{N}(0, 1), \ j = 1, \ldots, m, \ \text{i.i.d.} \quad \Rightarrow \quad \sum_{j=1}^{m} X_j^2 \sim \chi^2(m),$$

(10) Student distribution.
If $Z \sim \mathcal{N}(0, 1)$, $Y \sim \chi^2(m)$, $Z \perp Y$, then,

$$T := \frac{Z}{\sqrt{Y/m}},$$

has a student distribution with m degrees of freedom.
Its density is given by

$$f_T(t) = \frac{\Gamma((m+1)/2)}{\sqrt{m\pi} \ \Gamma(m/2)} \left(1 + \frac{t^2}{m}\right)^{-(m+1)/2}, \quad t \in \mathbb{R}.$$

(11) Studentizing. Let $\{X_i\}_{i=1}^n$ be i.i.d. with $\mathcal{N}(\mu, \sigma^2)$ distribution. Let $\overline{X}_n := \sum_{i=1}^n X_i/n$ and set

$$S_n^2 := \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2.$$

Then, $\overline{X}_n \perp S_n^2$ and

$$\frac{\sqrt{n} \left[\overline{X}_n - \mu\right]}{S_n}$$

has a Student distribution with $n-1$ degrees of freedom.

3. Borel-Cantelli

(12) Infinitely often, i.o.. For a given countable sequence of events $\{A_n\}_{n=1}^{\infty}$, the set $\{A_n \text{ i.o.}\}$ is defined by

$$\{A_n \text{ i.o.}\} := \cap_{n=1}^{\infty} \cup_{m=n}^{\infty} A_m.$$

(13) Borel-Cantelli Lemma 1. Suppose that $\{A_n\}_{n=1}^{\infty}$ satisfy

$$\sum_n P(A_n) < \infty.$$

Then, $P(\{A_n \text{ i.o.}\}) = 0$.

(14) Borel-Cantelli Lemma 2. Suppose that $\{A_n\}_{n=1}^{\infty}$ are mutually independent and satisfy

$$\sum_i P(A_n) = \infty.$$

Then, $P(\{A_n \text{ i.o.}\}) = 1$.

4. Limit theorems

(15) Law of Large Numbers. Let $\{X_i\}_{i=1}^{\infty}$ be an i.i.d. sequence. Set $\mu := EX_i$, $\sigma^2 := \mathrm{Var}(X_i) < \infty$ for any i, and

$$\overline{X}_n := \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Then the weak law of large numbers states that $\overline{X}_n$ converges to μ in probability and the strong law states that the convergence is almost surely.

(16) Central Limit Theorem. Let $\{X_i\}_{i=1}^{\infty}$ be an i.i.d. sequence. Set $\mu := EX_i$, $\sigma^2 := \mathrm{Var}(X_i) < \infty$ for any i, and $\overline{X}_n := \sum_{i=1}^{n} X_i/n$. Let

$$Z_n := \frac{\sqrt{n}}{\sigma} [\overline{X}_n - \mu].$$

The distribution of Z_n converges to the standard Gaussian, i.e., for all $z \in \mathbb{R}$

$$\lim_{n\to\infty} P(Z_n \le z) = \Phi(z),$$

where Φ is the standard normal distribution function.

5. Inequalities

(17) Jensen's Inequality. For a convex function $g : \mathbb{R} \to \mathbb{R}$,

$$g(EX) \le Eg(X).$$

(18) Markov's Inequality. For a non-negative random variable X and constant $a > 0$

$$P(X \ge a) \le \frac{EX}{a} .$$

Generalized Chebyshev's Inequality. For a random variable $X \in \mathbb{R}$ and a non-negative increasing function g and a real number a with $g(a) > 0$,

$$P(X \ge a) \le \frac{Eg(X)}{g(a)} .$$

6. Moments

(19) Variance and Standard Deviation.

$$\text{variance of X} = \mathrm{Var}(X) := EX^2 - (EX)^2 = E(X - EX)^2,$$
$$\text{standard deviation of X} = \sigma_X := \sqrt{\mathrm{Var}(X)}.$$

(20) Covariance and Correlation.

$$\begin{aligned}\text{covariance of X and Y} = \mathrm{Cov}(X,Y) &:= EXY - (EX)(EY)\\ &= E(X-EX)(Y-EY),\end{aligned}$$
$$\text{correlation of X and Y} = \mathrm{Corr}(X,Y) := \frac{\mathrm{Cov}(X,Y)}{\sigma_X\ \sigma_Y} \in [-1,1].$$

We have

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

(21) Moment Generating Function. The moment generating function of a random variable X is

$$\Psi(t) := \mathbb{E}\left[e^{tX}\right], \quad t \in \mathbb{R}.$$

The value of $\Psi(t)$ could be $+\infty$. When it is finite for t near the origin,

$$EX^k = \frac{d^k}{dt^k}\Psi(0), k = 1, 2, \ldots.$$

7. Confidence interval for the mean of a normal distribution

(22) Two sided Confidence Interval. Let $X_1, \ldots, X_n$ be i.i.d. $\mathcal{N}(\mu, \sigma^2)$ and

$$\overline{X}_n := \sum_{i=1}^n X_i/n,\ S_n^2 := \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2.$$

Set

$$A := \overline{X}_n - T_{n-1}^{-1}(1-\alpha/2)\ \frac{S_n}{\sqrt{n}},$$
$$B := \overline{X}_n + T_{n-1}^{-1}(1-\alpha/2)\ \frac{S_n}{\sqrt{n}},$$

where T_m is the c.d.f. of the Student distribution with m degrees of freedom, T_m^{-1} is its inverse function and $0 < \alpha < 1$. The interval (A, B) is a two sided $(1-\alpha)$-confidence interval for μ.

(23) One sided Confidence Interval. Let $X_1, \ldots, X_n$ be i.i.d. $\mathcal{N}(\mu, \sigma^2)$ and

$$\overline{X}_n := \sum_{i=1}^n X_i/n,\ S_n^2 := \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2.$$

Set

$$\underline{A} := \overline{X}_n - T_{n-1}^{-1}(1-\alpha)\ \frac{S_n}{\sqrt{n}},$$
$$\underline{B} := \overline{X}_n + T_{n-1}^{-1}(1-\alpha)\ \frac{S_n}{\sqrt{n}}.$$

where T_m is the c.d.f. of the Student distribution with m degrees of freedom, T_m^{-1} is its inverse function and $0 < \alpha < 1$.

The interval $(\underline{A}, \infty)$ is an upper $(1-\alpha)$-confidence interval for μ.

The interval $(-\infty, \underline{B})$ is a lower $(1-\alpha)$-confidence interval for μ.

8. Various

(24) Change of Variables Formula. Let $X = (X_1, \ldots, X_n)$ have a continuous joint distribution $f_X(x)$ for $x \in \mathbb{R}^n$. Let $Y = AX$ for some non-singular square matrix A. Then, the probability distribution function of Y is given by

$$f_Y(y) = \frac{1}{|\det(A)|}\, f_X(A^{-1}y), \quad y \in \mathbb{R}^n.$$

(25) Continuous Bayes Theorem. Consider two random variables X and θ, where θ has density $w(\cdot)$ and given $\theta = \vartheta$, the random variable X has density $f(x \mid \vartheta)$. Then the density of θ given $X = x$ is

$$w(\vartheta \mid x) = \frac{f(x \mid \vartheta)\; w(\vartheta)}{f(x)},$$

where

$$f(x) = \int f(x|\vartheta)\; w(\vartheta)\; d\vartheta.$$