Estatística Numérica e Computacional

Tutorial Notes

Marta Faias

2018/2019

FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA
Departamento de Matemática

First Part Syllabus

1. Elements of probability.
2. Pseudo-random number generation (discrete and continuous).
3. Monte Carlo methods.
   3.1 Variance reduction techniques.
4. Resampling techniques: Bootstrap and Jackknife.
5. Newton-Raphson method.

References

Alphabetic order

1. Davison, A.C., Hinkley, D.V., Bootstrap Methods and their Application, Cambridge University Press, 1997.
2. Gentle, J.E., Random Number Generation and Monte Carlo Methods, Springer-Verlag, 1998
3. Hossack, I.B., Pollard, J.H., Zehnwirth, B., Introductory Statistics with Applications in General Insurance, Cambridge University Press, 2nd Edition, 1999.
4. Ross, S.M., Simulation, 3rd Edition, Academic Press, 2002.

1 Elements of probability

$\Omega \equiv$ Space of possible outcomes of some experience.
$\mathcal{S} \equiv$ Set of events over Ω.

Definição: Let $(\Omega, \mathcal{S})$ a sample space. A **random variable** X (v.a) is a real function with finite values, $X : \Omega \to \mathbb{R}$ such that for each $x \in \mathbb{R}$

$$A_x = \{\omega \in \Omega \, : \, X(\omega) \leq x\} \subset \mathcal{S}$$

that is, A_x is an event.

Remark: If X is a random variable and $g : \mathbb{R} \to \mathbb{R}$ is a function then $Y = g(X)$ is also a random variable.

Let P denote a probability measure over the event set $\mathcal{S}$.

1 Elements of probability

Definição (**Distribution function**): The distribution function of the random variable X is:

$$F_X(x) = P(X \leq x) = P\{\omega : X(\omega) \leq x\}, \forall x \in R.$$

Proposition Let X be a randam variable with distribution funcion F then,

$$P(X = x) = P(X \leq x) - P(X < x) = F(x) - F(x^-), \quad \forall x \in \mathbb{R}$$

where $F(x^-) = \lim_{t \to x^-} F(t)$.

1 Elements of probability

$$D = \{a \in \mathbb{R} : p(X = a) > 0\}$$

Discrete random variable: A random variable is discrete if the set D is at most countable and $P(X \in D) = 1$.

Probability function: Let X be a discrete random variable. The probability function (f.p.) or probability mass function of X is the function that maps each element of D into its probability:

$$X = \begin{cases} x_1 & x_2 & \cdots & x_i & \cdots \\ p_1 = P(X = x_1) & p_2 = P(X = x_2) & \cdots & p_i = P(X = x_i) & \cdots \end{cases}$$

Properties:

- $P(X = x_i) = f(x_i) = p_i \geq 0$;
- $\sum_{i=1}^{\infty} p_i = 1$.

1 Elements of probability

Continuous random variable: A random variable X is continuous if $D = \emptyset$ and there exists a non-negative function f such that for for $I \subset \mathbb{R}$,

$$P(X \in I) = \int_I f(x)dx.$$

The function f is called **density function**.

Properties:

- $f(x) \geq 0, \quad \forall x \in \mathbb{R};$
- $\int_{-\infty}^{+\infty} f(x)dx = 1.$

1 Elements of probability - Moments

Mean of random variable:

$$\mu = E(X) = \begin{cases} \sum_{i=1}^{\infty} x_i P(x = x_i) & \text{if } X \text{ is a discrete r.v.} \\ \\ \int_{-\infty}^{+\infty} x f(x) dx & \text{if } X \text{ is a continuous r.v.} \end{cases}$$

Properties:

- If a is a constant $E(a) = a$;
- If a and b are constants, $E(aX + b) = aE(X) + b$;
- If there exist $E(g_1(X))$ and $E(g_2(X))$ then $E(g_1(X) + g_2(X)) = E(g_1(X)) + E(g_2(X))$.

Variance:

$$\sigma^2 = Var(X) = E((X - E(X))^2) = E(X^2) - (E(X))^2.$$

Or,

$$\sigma^2 = Var(X) = E((X - \mu))^2) = E(X^2) - \mu^2.$$

Standard deviation:

$$\sigma = \sqrt{Var(X)}.$$

Properties:

- If a is a constant $Var(a) = 0$;
- If a and b are constants, $Var(aX + b) = a^2 Var(X)$;

1 Elements of probability

1.1 Discrete random variables

Binomial: $X \sim B(n, p)$ represents the number of successes that occur in n independent trials when p is the probability of success.

f.m.p: $p_i = P\{X = i\} = \binom{n}{i} p^i (1-p)^{n-i}, \; i = 0, 1, ..., n.$

$\binom{n}{i} = \frac{n!}{i!(n-i)!)}.$

$$E[X] = np, \qquad Var(X) = np(1-p)$$

Bernoulli: $X \sim Ben(1, p)$ represents a binomial random variable with parameters $(1, p)$

$$E[X] = p, \qquad Var(X) = p(1-p)$$

Geometric: $X \sim Geo(p)$ represents the number of the first trial that is a success in a sequence of independent trials when p is the probability of success..

f.m.p: $P\{X = n\} = p(1-p)^{n-1}, \; n \geq 1.$

$$E[X] = \frac{1}{p}, \qquad Var(X) = \frac{1-p}{p^2}.$$

Poisson: $X \sim Poi(\lambda)$ represents the number of successes in an infinite sequence of experiences, $X = 0, 1, 2, ...$

f.m.p: $p_i = P\{X = i\} = e^{-\lambda}\frac{\lambda^i}{i!}, \; i = 0, 1, ...$

$$E[X] = Var(X) = \lambda$$

The negative Binomial random variable: $X \sim NB(p, r)$ represents the number of trials need to obtain a total of r successes when each trial is independently a sucess with probability p.

f.m.p: $P\{X = n\} = \binom{n-1}{r-1} p^r (1-p)^{n-r}, \ n \geq r.$

$$E[X] = \frac{r}{p}, \qquad Var(X) = \frac{r(1-p)}{p^2}.$$

Hypergeometric random variable: $X \sim Hyp(M, N, n)$ Consider a box with $N + M$ balls, N are white colored and M are black colored. If we choose a random sample with n balls, then X represents the total of white balls that we have chosen.

f.m.p:

$$P\{X = i\} = \frac{\binom{N}{i}\binom{M}{n-i}}{\binom{N+M}{n}}, \quad \max(0, M+n-N) \leq i \leq \min(M, n).$$

$$E[X] = \frac{nN}{N+M}, \qquad Var(X) = Var(X) = \frac{nNM}{(N+M)^2}\left(1 - \frac{n-1}{N+M-1}\right).$$

1.2 Continuous random variables

Uniforme: $U \sim U(a, b)$

f.d.p: $$f(x) = \begin{cases} \dfrac{1}{b-a} & se \quad a < x < b \\ 0 & se \quad x \leq a \text{ ou } x \geq b \end{cases}.$$

Distribution function: $F(x) = P\{X \leq x\} = \dfrac{x-a}{b-a}$

$$E[X] = \frac{b+a}{2}, \qquad Var(X) = \frac{1}{12}(b-a)^2$$

Exponencial: $X \sim Exp(\lambda)$.
f.d.p: $f(x) = \lambda e^{-\lambda x}$, $0 < x < \infty$

Distribution function: $F(x) = 1 - e^{-\lambda x}$.

$$E[X] = \frac{1}{\lambda}, \qquad Var(X) = \frac{1}{\lambda^2}$$

$P(X > t + s \,|\, X > s) = p(X > t)$ (without memory)

Normal: $X \sim N(\mu, \sigma^2)$.

f.d.p: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \ -\infty < x < \infty$

Distribution function:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx, \ -\infty < x < \infty$$

$$E[X] = \mu, \qquad Var(X) = \sigma^2.$$

Gamma: $X \sim Gama(\alpha, \beta)$.

f.d.p: $f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta x} x^{\alpha-1}, \ x > 0, \quad \Gamma(\alpha) = \int_0^{+\infty} e^{-x} x^{\alpha-1} dx.$

Distribution function: $\Phi(x) = \int_0^x \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta x} x^{\alpha-1} dx, \ x > 0.$

$$E[X] = \frac{\alpha}{\beta}, \qquad Var(X) = \frac{\alpha}{\beta^2}.$$

1.3 Some probability results

Markov's Inequality: If X is a random variable which takes only positive values, then for any value $a > 0$

$$P\{X \geq a\} \leq \frac{E[X]}{a}.$$

Chebyshev's Inequality: If X is a random variable having mean μ and variance σ^2, then for any value $k > 0$,

$$P\{|X - \mu| \geq k\sigma\} \leq \frac{1}{k^2}$$

.

The weak law of large numbers: Let $X_1, X_2, ..., X_n$ be a sequence of independent and identically distributed random variables having mean μ. Then, for any $\epsilon > 0$,

$$P\left\{\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| > \epsilon\right\} \to 0 \text{ when } n \to \infty.$$

The strong law of large numbers: Let $X_1, X_2, ..., X_n$ be a sequence of independent and identically distributed random variables having mean μ. Then,

$$P\left\{\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu\right\} = 1.$$

The central limit theorem

Let $X_1, X_2, ..., X_n$ be a sequence of independent and identically distributed random variables having mean μ and finite variance σ^2. Then,

$$\lim_{n\to\infty} P\left\{\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} < x\right\} = \Phi(x).$$

$\longrightarrow$ Φ denotes the distribution function of a standard normal random variable.

Or, if $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_n$,

$$Z_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \overset{a}{\sim} N(0,1).$$

2 Pseudo-random number generation (discrete and continuous)

2.1 Discrete random variables

A procedure to generate the value of a discrete variable X with finite support

$$\begin{array}{c|cccc} X & x_0 & x_1 & \cdots & x_n \\ \hline P(X = x_j) & p_0 & p_1 & \cdots & p_n \end{array},$$

that is,

$$P\{X = x_j\} = p_j,\ j = 0, 1, 2, ..., n.$$

We generate a random number U, $U \sim U(0,1)$ and then we run the following procedure

$$X = \begin{cases} x_0 & se \quad U < p_0 \\ x_1 & se \quad p_0 \leq U < p_0 + p_1 \\ \vdots & \\ x_j & se \quad \sum_{i=0}^{j-1} p_i \leq U < \sum_{i=0}^{j} p_i \\ \vdots & \end{cases}$$

Actually, $P(x = x_j) = P\left\{ \sum_{i=0}^{j-1} p_i \leq U < \sum_{0=1}^{j} p_i \right\} = p_j$, that is,

$X = x_j$ if $F(x_{j-1}) \leq U < F(x_j)$, where F is the distribution function of X.

The above procedure can be implemented running the following algorithm

Algorithm:
If $U < p_0$, set $X = x_0$ and stop;
If $U < p_0 + p_1$, set $X = x_1$ and stop;
if $U < p_0 + p_1 + p_2$, set $X = x_2$ and stop;...

In practice we generate a random number U and after we choose the value of X by finding the interval $[F(x_{j-1}), F(x_j))$ where U belongs.

A discrete uniform random variable

In the case of a discrete uniform random variable,

X	1	2	$\cdots$	n
$P(X = x_j)$	$\frac{1}{n}$	$\frac{1}{n}$	$\cdots$	$\frac{1}{n}$

that is,

$$P\{X = j\} = \frac{1}{n}, \; j = 1, 2, ..., n$$

the previous algorithm is simplified,

$$P\{X = j\} = \frac{1}{n}, \; j = 1, 2, ..., n$$

$$X = j \text{ se } \frac{j-1}{n} \leq U < \frac{j}{n} \Leftrightarrow X = \text{Int}(nU) + 1.$$

The Poisson random variable

$p_i = P\{X = i\} = e^{-\lambda}\frac{\lambda^i}{i!}, \ i = 0, 1, 2, ...$

For this random variable there is a recursion between p_{i+1} and p_i,

$$p_{i+1} = \frac{\lambda}{1+i} p_i, \ \ i \geq 0.$$

This allows to simplify the algorithm to nd then we the algorithm to generate this variable can

Algorithm:

1. Generate a random number U;
2. $i = 0$, $p = e^{-\lambda}$, $F = p$;
3. If $U < F$, set $X = i$ and stop;
4. else, $p = \frac{\lambda}{i+1} p$, $F = F + p$, $i = i + 1$;
5. Return to step 3.

The Binomial random variable

$$p_i = P\{X = i\} = \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i},\ i = 0, 1, .., n$$

For this random variable there is a recursion between p_{i+1} and p_i,

$$p_{i+1} = \frac{n-i}{i+1} \frac{p}{1-p} p_i,\ \ i \geq 0.$$

Algoritmo:

1. Generate a random number U;
2. $c = \dfrac{p}{1-p}$, $i = 0$, $pr = (1-p)^n$, $F = pr$;
3. If $U < F$, set $X = i$ and stop;
4. else, $pr = [c\dfrac{(n-i)}{i+1}]pr$, $F = F + pr$, $i = i + 1$;
5. Return to step 3.

The acceptance-rejection technique

Let X and Y be two random variables with the same support D but with different probability mass functions.

- $\{p_j, j \geq 0\}$ is the probability function of X;
- $\{q_j, j \geq 0\}$ is the probability function of Y;

Suppose that it is **easy** to simulate the random variable Y and our aim is to generate the random variable X.
Suppose that there exists $c > 0$ such that,

$$\frac{p_j}{q_j} \leq c, \ \ \forall j : p_j > 0$$

Algoritm:

1. Generate a random variable Y, with probability mass function q_Y.
2. Generate a random number U.
3. If $U \leq \frac{p_Y}{cq_Y}$, set $X = Y$ and stop, else, return to step 1.

Actually, $P\{X = j\} = p_j$, $j = 0, 1, \ldots$

Proposition: The Inverse Transformation Method.
Let U be a random variable with uniform distribution over the interval $(0,1)$. For any continuous distribution function F, the distribution function for the random variable $X = F^{-1}(U)$ is the function F.

Proof: Let $X = F^{-1}(U)$,

$$P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$$

then,

$$P(X \leq x)) = F(x).$$

Remark: When $U \sim U(0,1)$, $P(U \leq u) = u$.

Example.

Generate an exponential random variable with parameter λ using the inverse tranformation method,

Since $F(x) = 1 - e^{-\lambda x}$,

then, $X = F^{-1}(U) = -\frac{1}{\lambda} log(U)$ is an exponential random variable with parameter λ.

Example. The acceptance-rejection technique can also be applied to continuous random variables.
Let us generate a normal standard random variable.
$X \sim N(0,1)$, the density function for $|X|$:

$$f(x) = \frac{2}{\sqrt{2\pi}} e^{-x^2/2},\ 0 < x < \infty$$

Let $g(x) = e^{-x}$, $0 < x < \infty$

$\dfrac{f(x)}{g(x)} = \sqrt{2/\pi}\, e^{x - x^2/2}$, then $c = \mathrm{Max}\dfrac{f(x)}{g(x)} = \dfrac{f(1)}{g(1)} = \sqrt{2e/\pi}$.

Thus,

$$\frac{f(x)}{cg(x)} = exp\Big\{ -\frac{(x-1)^2}{2} \Big\}$$

Algoritm:

1. Generate an exponential random variable Y with parameter $\lambda = 1$;
2. Generate a random number U;
3. If $U \leq exp\Big\{-\dfrac{(Y-1)^2}{2}\Big\}$, set $X = Y$;
4. else, return to step 1.