Convergence concepts

Delta method: proof

Theorem (Delta Method): If $\sqrt{n}(Y_n - m) \xrightarrow{d} N(0, c^2)$ and $g'(m) \neq 0$, then

$$\sqrt{n}[g(Y_n) - g(m)] \xrightarrow{d} N(0, [g'(m)]^2 c^2).$$

Proof:

1. Since $\underbrace{\sqrt{n}(Y_n - m) \xrightarrow{d} N(0, c^2)}_{\text{CLT (Based on the assumption)}}$ and $\underbrace{Z_n = \frac{1}{\sqrt{n}} \xrightarrow{p} 0}$, by Slutsky's theorem

$$(Y_n - m) = Z_n \cdot \sqrt{n}(Y_n - m) \xrightarrow{d} 0 \quad \text{as } n \to \infty$$

$\Rightarrow Y_n - m \xrightarrow{d} 0$ as $n \to \infty$

$\Leftrightarrow Y_n - m \xrightarrow{P} 0$ as $n \to \infty$

$\Leftrightarrow Y_n \xrightarrow{P} m$ (I)

2. Define a function $u(y)$ which is continuous at $x = m$ by (II)

$$u(y) = \begin{cases} \frac{g(y)-g(m)}{y-m} - g'(m) & y \neq m \\ 0 & y = m \end{cases}$$

$$\lim_{y \to m} u(y) = g'(m) - g'(m) = 0 = u(m)$$

Note:
$$u(Y_n) = \frac{g(Y_n) - g(m)}{Y_n - m} - g'(m)$$
$$(Y_n - m)\left[g'(m) + u(Y_n)\right] = g(Y_n) - g(m)$$

3. Note that:

Since $u(y)$ is continuous at $y = m$ + $Y_n \xrightarrow{P} m$

$\Rightarrow u(Y_n) \xrightarrow{P} u(m) = 0$ as $n \to \infty$

4. Write

$$\sqrt{n}[g(Y_n) - g(m)] = \underbrace{\sqrt{n}(Y_n - m)}_{\xrightarrow{d} N(0,c^2)} \underbrace{\boxed{\left[g'(m) + u(Y_n)\right]}}_{g'(m)} \xrightarrow{P} g'(m)$$

$u(Y_n) \xrightarrow{P} 0$

Slutsky's theorem

$$\sqrt{n}\left[g(Y_n) - g(m)\right] \xrightarrow{d} \underbrace{g'(m) N(0, C^2)}_{\sim N(0, (g'(m))^2 C^2)}$$

① Practice Problem for Convergence in dist/prob/delta method/CLT will be Uploaded Soon (By tonight)

② Bring two-pages formula sheets (Sheet1 Sheet 2)

③ I will ask all the topics from lecture 27

$\longrightarrow$ Conditional dist.

④ I will provid the table of dist/mean/varince/MGF

Convergence concepts

Delta method: a heuristic

Normal approximation obtained by delta method has a heuristic justification:

- Since $\sqrt{n}(Y_n - m) \xrightarrow{d} N(0, c^2)$, we expect Y_n to be close to m for large n

for large n, Y_n is close to m,

- Do a Taylor series expansion of $g(Y_n)$ around m (ignore higher order terms)

$$g(Y_n) \approx g(m) + g'(m)(Y_n - m)$$

$$\sqrt{n}\left(g(Y_n) - g(m)\right) \dot{\approx} g'(m)\sqrt{n}(Y_n - m)$$

approximate

- If Y_n is asymptotically normal then $g(Y_n)$ will be too

Example: Suppose $X_1, X_2, \ldots$ are iid Poisson(λ)

$$\sqrt{n}\left(g(Y_n) - g(m)\right) \dot{\approx} g'(m)\sqrt{n}(Y_n - m)$$

Let $Y_n = g(\bar{X}_n) = 2\sqrt{\bar{X}_n}$ $\quad g(x) = 2\sqrt{x}$, $\quad g'(x) = \frac{1}{\sqrt{x}}$ $\quad g'(m) = \frac{1}{\sqrt{m}}$

$$\sqrt{n}\left(2\sqrt{\bar{X}_n} - 2\sqrt{m}\right) \approx \frac{1}{\sqrt{m}}\sqrt{n}(\bar{X}_n - m)$$

Note:

$m = \mathbb{E}[\bar{X}_n] = \mathbb{E}[X] \xrightarrow{X \sim \text{Poisson}(\lambda)} \lambda$

$$\sqrt{n}\left(2\sqrt{\bar{X}_n} - 2\sqrt{\lambda}\right) \approx \frac{1}{\sqrt{\lambda}}\underbrace{\sqrt{n}(\bar{X}_n - \lambda)}$$

$$\frac{1}{\sqrt{\lambda}} N(0, \lambda^2) \sim N\left(0, \frac{\lambda^2}{\lambda}\right) = N(0, \lambda)$$

STAT 542: Summary to date
Where we have been
& where we are headed

- Completed
  - Probability and random variables (definition, cdf, pdf/pmf)
  - Univariate distributions
    (definitions, expectation, transformations, families)
  - Multivariate distributions
    (joint distribution, covariance, conditional distribution, marginal distribution, independence, transformations, order statistics)
  - Convergence concepts (e.g., convergence in distribution or in probability, CLT, WLLN, Delta Method)

- Next
  - Introduction to stochastic/probabilistic simulation
  - skip Introduction to stochastic processes: Poisson processes, standard Brownian motion, discrete space Markov chains

Probabilistic Simulation

Introduction

- Section 5.6 of Casella & Berger is a basic introduction to stochastic/probabilistic simulation.

- The idea is that if I have some complicated probability model, rather than trying to do calculus or numerical analysis, in order to compute quantities of interest (e.g., probabilities/expected values), it may simply be easier to simulate a large number of realizations from the model and look at their characteristics.

- Topics

  1. Uniform Number Generators

  2. General Methods for Simulation: Discrete and Continuous Distributions

  3. Simulation Tricks for Standard Distributions

  4. Rejection Sampling Algorithm

  5. Importance Sampling

Probabilistic Simulation

Uniform$(0, 1)$ Generation

Basis of all stochastic simulation is generating values that look like *iid* $U(0, 1)$ *realizations*

- These can be "physical random numbers" obtained from "physical random processes" like radioactive decay (transformed inter-event times for particle emissions)

- by far the most common method is the use of pseudo-random generators that are just recursive numerical algorithms:

  One popular method is the *congruential method*

  – for integers a, c, m, define

  $$x_i = (ax_{i-1} + c) \bmod m, \quad i = 1, 2, 3, \ldots,$$

  (i.e., $x_i < m$ is the integer remainder after substracting the maximum integer multiple of m from $ax_{i-1} + c$)

  Find integer q

  $(q+1)m > ax_{i+1} + c \geq qm$

  $x_i = (ax_{i-1}+c) \bmod m$

  – x_0 is some appropriately chosen seed

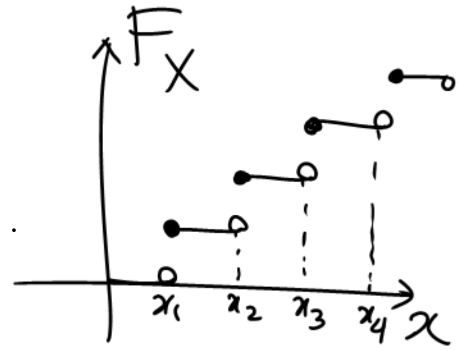
  – Define $u_i = x_i/m$, $i \geq 1$

  – Then, the stream of numbers $u_1, u_2, \ldots$ can "look like" (approximate) realizations of independent $U(0, 1)$ r.v.'s

  – e.g., IMSL Library uses $c = 0$, $a = 16807$, $m = 2^{31} - 1$

Probabilistic Simulation

General Method for Direct Simulation



General Method for Discrete Distributions

✓ 1. Suppose F is a discrete cdf with jumps at $x_1 < x_2 < x_3 < \cdots$.

2. Then for $U \sim U(0,1)$, define

$$\longrightarrow X = \begin{cases} x_1 & \text{if } U < F(x_1) \\ x_i & \text{if } F(x_{i-1}) \le U < F(x_i) \text{ for } i > 1 \end{cases}$$

3. Then, X is a r.v. with cdf F

for $i>1$

$$\mathbb{P}(X = x_i) = \mathbb{P}(F(x_{i-1}) < \underset{\substack{\downarrow \\ U \sim Uni(0,1)}}{U} < F(x_i)) = \int_{F(x_{i-1})}^{F(x_i)} 1 \, dy = F(x_i) - F(x_{i-1})$$

$$F_X(x_i) = \sum_{i=2}^{x_i} \mathbb{P}(X = x_i) = \sum_{i=2}^{x_i} F(x_i) - F(x_{i-1})$$

General Method for Continuous Distributions

1. Suppose F is a continuous cdf and, for $0 < p \le 1$, define

$$\longrightarrow F^{-1}(p) = \inf\{x \in \mathbb{R} : F(x) = p\}$$

2. If $U \sim U(0,1)$ then $X = F^{-1}(U)$ is a random variable with cdf F.

$$F_X(x) \underset{\text{def}}{=} \mathbb{P}(X \le x) = \mathbb{P}(F^{-1}(U) \le x)$$
$$= \mathbb{P}(F(F^{-1}(U)) \le F(x))$$
$$= \mathbb{P}(U \le F(x)) = \int_0^{F(x)} dy = F(x)$$

These general methods are fine in theory, but could be numerically challenging or inefficient to implement: there are some special tricks for simulating from standard distributions (illustrated next)