Statistical Methods MATH2715 info

Teaching material is all online!

- On Minerva `http://minerva.leeds.ac.uk`
- On GitHub `https://github.com/luisacutillo78/Statistical-Methods-Lecture-Notes`

R code submission

- No technichal issue - please submit your SURNAMEstudentid.R [or .Rmd as required] file in the assignment folder.
- **Please print a copy of your notebook and put it into your marker collection box**.

Resources

- Mathematical Statistics and Data Analysis - 3rd ed. (by J. A. Rice);
- http://www1.maths.leeds.ac.uk/statistics/R/Rintro.pdf;
- https://www.datacamp.com/courses/free-introduction-to-r.

Where We've Been, Where We're Going

In the previous Lecture

- Univariate and Multivariate Change of Variables
- Markov's Inequality
- Chebyshev's Inequality
- Weak Law of Large Numbers
- Exercises & Questions

Today

- Weak Law of Large Numbers: Interpretation and discussion
- Convergence in probability
- Convergence in distribution
- Central limit theorem

Why are we studying limit theorems?

Questions we are addressing

- What happens when we consider a long sequence of random variables ?
- What can we reasonably infer from data?
- Laws of large numbers: averages of random variables converge on expected value?
- Central Limit Theorems: sum of random variables have normal distribution?

Sequence of Random Variables

Sequence of Independent and Identically, Distributed Random variables.

- Sequence: $X_1, X_2, \ldots, X_n, \ldots$
- Think of a sequence as sampled data:
  - Suppose we are drawing a sample of N observations
  - Each observation will be a random variable, say X_i
  - With realization x_i

Mean/Variance of Sample Mean

Sample Mean

Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with mean μ and variance σ^2. Let $\bar{X}_n$ be the sample mean. Then $E[\bar{X}_n] = \mu$ and $\text{var}(\bar{X}_n) = \frac{\sigma^2}{n}$

Proof.

$$E[\bar{X}_n] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n} n\mu = \mu$$

$$\text{var}(\bar{X}_n) = \frac{1}{n^2}\text{var}(\sum_{i=1}^{n} X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\text{var}(X_i) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

$\square$

Weak Law of Large Numbers

Proposition

Suppose $X_1, X_2, \ldots, X_n$ is a random sample from a distribution with mean μ and $Var(X_i) = \sigma^2$. Then, for all $\epsilon > 0$,

$$P\left\{\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \mu\right| \geq \epsilon\right\} \to 0 \text{ as } n \to \infty$$

Weak Law of Large Numbers

Interpreation

It is a common belief that if we toss a coin textitmany times, the propostion of heads will be close to $frac12$.
The law of large numbers is a mathematical interpretation of this believe!

Example

- Successive tosses of a coin can be modelled as independent random trials X_i
- Each X_i takes on 0 (if the $i - thresultistail)or1(ifthe$i-th result is head)
- The proportion of heads in n trials is $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

The law of large numbers says that $\bar{X}$ *approaches* μ as the number of trials grows.

Sequences and Convergence: Recalls

Sequence of real numbers:

$$\{a_i\}_{i=1}^{\infty} = \{a_1, a_2, a_3, \ldots, a_n, \ldots, \}$$

Definition

We say that the sequence $\{a_i\}_{i=1}^{\infty}$ converges to real number A if for each $\epsilon > 0$ there is a positive integer N such that for $n \geq N$, $|a_n - A| < \epsilon$

Sequences and Convergence

Sequence of functions:

$$\{f_i\}_{i=1}^{\infty} \quad = \quad \{f_1, f_2, f_3, \ldots, f_n, \ldots,\}$$

Definition

Suppose $f_i : X \to \Re$ for all i. Then $\{f_i\}_{i=1}^{\infty}$ converges pointwise to f if, for all $x \in X$ and $\epsilon > 0$, there is an N such that for all $n \geq N$,

$$|f_n(x) - f(x)| < \epsilon$$

This is as strong of a statement!

Convergence: Some Definitions

Let $\widehat{\theta}_i$ be an estimator for θ based on i observations.

Convergence: Some Definitions

Let $\widehat{\theta}_i$ be an estimator for θ based on i observations.
Increasing sample size we get a Sequence of estimators:

$$\left\{\widehat{\theta}_i\right\}_{i=1}^{n} \quad = \quad \left\{\widehat{\theta}_1, \widehat{\theta}_2, \widehat{\theta}_3, \ldots, \widehat{\theta}_n\right\}$$

Question: What can we say about $\left\{\widehat{\theta}_i\right\}_{i=1}^{n}$ as $n \to \infty$?

- What is the probability $\widehat{\theta}_n$ differs from θ?

Convergence: Some Definitions

Let $\widehat{\theta}_i$ be an estimator for θ based on i observations.
Increasing sample size we get a Sequence of estimators:

$$\left\{\widehat{\theta}_i\right\}_{i=1}^{n} = \left\{\widehat{\theta}_1, \widehat{\theta}_2, \widehat{\theta}_3, \ldots, \widehat{\theta}_n\right\}$$

Question: What can we say about $\left\{\widehat{\theta}_i\right\}_{i=1}^{n}$ as $n \to \infty$?

- What is the probability $\widehat{\theta}_n$ differs from θ?
- What is the probability $\left\{\widehat{\theta}_i\right\}_{i=1}^{n}$ converges to θ?

Convergence: Some Definitions

Let $\widehat{\theta}_i$ be an estimator for θ based on i observations.
Increasing sample size we get a Sequence of estimators:

$$\left\{\widehat{\theta}_i\right\}_{i=1}^{n} = \left\{\widehat{\theta}_1, \widehat{\theta}_2, \widehat{\theta}_3, \ldots, \widehat{\theta}_n\right\}$$

Question: What can we say about $\left\{\widehat{\theta}_i\right\}_{i=1}^{n}$ as $n \to \infty$?

- What is the probability $\widehat{\theta}_n$ differs from θ?
- What is the probability $\left\{\widehat{\theta}_i\right\}_{i=1}^{n}$ converges to θ?

Convergence in Probability

Definition

We will say the sequence $\widehat{\theta}_n$ converges in probability to θ (perhaps a non-degenerate RV) if,

$$\lim_{n\to\infty} Prob(|\widehat{\theta}_n - \theta| > \epsilon) = 0$$

For any $\epsilon > 0$

Convergence in Probability

Definition

We will say the sequence $\widehat{\theta}_n$ converges in probability to θ (perhaps a non-degenerate RV) if,

$$\lim_{n\to\infty} Prob(|\widehat{\theta}_n - \theta| > \epsilon) = 0$$

For any $\epsilon > 0$

- ϵ is a tolerance parameter: how much error around θ?

Convergence in Probability

Definition

We will say the sequence $\widehat{\theta}_n$ converges in probability to θ (perhaps a non-degenerate RV) if,

$$\lim_{n\to\infty} Prob(|\widehat{\theta}_n - \theta| > \epsilon) = 0$$

For any $\epsilon > 0$

- ϵ is a tolerance parameter: how much error around θ?
- In the limit, convergence in probability implies the $\widehat{\theta}_n$ distribution collapses on a spike at θ

Convergence in Probability

Definition

We will say the sequence $\widehat{\theta}_n$ converges in probability to θ (perhaps a non-degenerate RV) if,

$$\lim_{n\to\infty} Prob(|\widehat{\theta}_n - \theta| > \epsilon) = 0$$

For any $\epsilon > 0$

- ϵ is a tolerance parameter: how much error around θ?
- In the limit, convergence in probability implies the $\widehat{\theta}_n$ distribution collapses on a spike at θ
- $\left\{\widehat{\theta}_i\right\}$ does not need actually converge to θ, only P($|\theta_n$ - $\theta| > \epsilon) = 0$

Convergence in Distribution

Definition

$\widehat{\theta}_n$, with cdf $F_n(x)$, converges in distribution to random variable Y with cdf $F(x)$ if

$$\lim_{n\to\infty} |F_n(x) - F(x)| = 0$$

For all $x \in \Re$ where $F(x)$ is continuous.

- Says that cdfs are equal, says nothing about convergence of underlying RV

Convergence in Distribution

Definition

$\widehat{\theta}_n$, with cdf $F_n(x)$, converges in distribution to random variable Y with cdf $F(x)$ if

$$\lim_{n \to \infty} |F_n(x) - F(x)| = 0$$

For all $x \in \Re$ where $F(x)$ is continuous.

- Says that cdfs are equal, says nothing about convergence of underlying RV
- Useful for justifying use of some sampling distributions

Central Limit Theorem

Proposition

Let X_1, $X_2, \ldots$ be a sequence of independent random variables with mean μ and variance σ^2. Let X_i have a cdf $P(X_i \leq x) = F(x)$ and moment generating function $M(t) = E[e^{tX_i}]$. Let $S_n = \sum_{i=1}^{n} X_i$. Then

$$\lim_{n\to\infty} P\left(\frac{S_n - \mu n}{\sigma\sqrt{n}} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{x^2}{2}\right)$$

Central Limit Theorem

Proposition

Let X_1, $X_2, \ldots$ be a sequence of independent random variables with mean μ and variance σ^2. Let X_i have a cdf $P(X_i \leq x) = F(x)$ and moment generating function $M(t) = E[e^{tX_i}]$. Let $S_n = \sum_{i=1}^{n} X_i$. Then

$$\lim_{n\to\infty} P\left(\frac{S_n - \mu n}{\sigma\sqrt{n}} \leq x\right) \quad = \quad \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{x^2}{2}\right)$$

Proof plan:

1) Rely on Fact that convergence of MGFs$\rightsquigarrow$ convergence in CDFs
2) Show that MGFs, in limit, converge on normal MGF

Proposition

Let F_n be a sequence of cumulative distribution functions with the corresponding moment generating functions M_n. F be a cdf with the moment generating functions M. If $\lim_{n\to\infty} M_n(t) \to M(t)$ for all t in some interval, then $F_n(x) \rightsquigarrow F(x)$ for all x (when F is continuous).

Proposition

Suppose $\lim_{n\to\infty} a_n \to a$, then

$$\lim_{n\to\infty}\left(1+\frac{a_n}{n}\right)^n = e^a$$

Proposition

Suppose $M(t)$ is a moment generating function some random variable X. Then $M(0) = 1$.

Proof of Central Limit Theorem

Proof. Suppose $X_1, \ldots, X_n$ are iid variables with $E[X] = 0$, variance σ_x^2, Moment Generating Function (MGF) $M_x(t)$.

Proof of Central Limit Theorem

Proof. Suppose $X_1, \ldots, X_n$ are iid variables with $E[X] = 0$, variance σ_x^2, Moment Generating Function (MGF) $M_x(t)$.
Let $S_n = \sum_{i=1}^{n} X_i$ and $Z_n = \frac{S_n}{\sigma_x \sqrt{n}}$.

Proof of Central Limit Theorem

Proof. Suppose $X_1, \ldots, X_n$ are iid variables with $E[X] = 0$, variance σ_x^2, Moment Generating Function (MGF) $M_x(t)$.
Let $S_n = \sum_{i=1}^{n} X_i$ and $Z_n = \frac{S_n}{\sigma_x \sqrt{n}}$.
$M_{S_n} = (M_x(t))^n$ and $M_{Z_n}(t) = \left(M_x \left(\frac{t}{\sigma_x \sqrt{n}} \right) \right)^n$

Proof of Central Limit Theorem

Proof. Suppose $X_1, \ldots, X_n$ are iid variables with $E[X] = 0$, variance σ_x^2, Moment Generating Function (MGF) $M_x(t)$.
Let $S_n = \sum_{i=1}^n X_i$ and $Z_n = \frac{S_n}{\sigma_x\sqrt{n}}$.
$M_{S_n} = (M_x(t))^n$ and $M_{Z_n}(t) = \left(M_x\left(\frac{t}{\sigma_x\sqrt{n}}\right)\right)^n$
Using Taylor's Theorem we can write

$$M_x(s) = M_x(0) + sM_x^{'}(0) + \frac{1}{2}s^2M_x^{''}(0) + e_s$$

$e_s/s^2 \to 0$ as $s \to 0$.

$$M_x(s) \;=\; M_x(0) + sM_x^{'}(0) + \frac{1}{2}s^2 M_x^{''}(0) + e_s$$

Filling in the values we have

$$M_x(s) \;=\; 1 + 0 + \frac{\sigma_x^2}{2}s^2 + \underbrace{e_s}_{\text{Goes to zero}}$$

Set $s = \frac{t}{\sigma_x \sqrt{n}}$ $\lim_{n\to\infty} s \to 0$. Then

$$\begin{aligned}
M_{Z_n}(t) &= \left(1 + \frac{\sigma_x^2}{2}\left(\frac{t}{\sigma_x\sqrt{n}}\right)^2\right)^n \\
&= \left(1 + \frac{t^2/2}{n}\right)^n \\
\lim_{n\to\infty} M_{Z_n}(t) &= e^{\frac{t^2}{2}}
\end{aligned}$$

WHITE BOARD EXERCISES

Today

- Review
- Questions