

高等统计学——第三周

Chapter 2

Chapter 2

Preliminaries

Some useful results in probability

Order relations

Chapter 3

Central Limit Theorems and Related Refinements

3.1.1 CLT for i.i.d. r.v.s

3.1.2 CLT for triangular arrays with finite variances

3.2 Uniform Berry-Esseen Bounds

3.3 Non-Uniform Berry-Essen Bounds

Preliminaries

Some useful results in probability

Let $\{X_n\}, \{Y_n\}, X, Y$ be r.v.'s on the same probability space (Ω, \mathcal{A}, P) , unless stated otherwise.

Theorem 2.1.1 (Algebraic operations)

1. If $X_n \rightarrow_p X$ and $Y_n \rightarrow_p Y$, then $X_n \pm Y_n \rightarrow_p X \pm Y$
2. If $X_n \rightarrow X$ a.s. and $Y_n \rightarrow Y$ a.s., then $X_n \pm Y_n \rightarrow X \pm Y$ a.s.
3. If $X_n \rightarrow_r X$ and $Y_n \rightarrow_r Y$, then $X_n \pm Y_n \rightarrow_r X \pm Y$
4. However, suppose that $X_n \rightarrow_d X$ and $Y_n \rightarrow_d Y$, where all r.v.'s may be defined on different probability spaces.

Then it is not true in general that $X_n \pm Y_n \rightarrow_d X \pm Y$

这个定理说明：依概率收敛、几乎处处收敛、r阶矩收敛都对加减法封闭，但是依分布收敛对加减法不一定封闭

定理第一条的证明思路：

$$\begin{matrix} X_n \rightarrow_p X \\ Y_n \rightarrow_p Y \end{matrix} \Rightarrow X_n \pm Y_n \rightarrow_p X \pm Y$$

$$P\{|X_n + Y_n - (X + Y)| > \epsilon\} \rightarrow 0$$

$$|(X_n - X) + (Y_n - Y)| > \epsilon$$

$$\{|(X_n - X) + (Y_n - Y)| > \epsilon\} \subset \{|X_n - X| > \frac{\epsilon}{2}\} \cup \{|Y_n - Y| > \frac{\epsilon}{2}\}$$

$$P\{|(X_n - X) + (Y_n - Y)| > \epsilon\} \subset P\{|X_n - X| > \frac{\epsilon}{2}\} \cup P\{|Y_n - Y| > \frac{\epsilon}{2}\}$$

$$\dots \rightarrow 0$$

定理最后一条的反例：

$$\delta \sim N(0, 1)$$

$$X_n \rightarrow_d \delta \quad Y_n \rightarrow_d -\delta$$

令 X_n 与 Y_n 独立同分布

构造 $Y_n \rightarrow_d \delta$

则 $X_n + Y_n \rightarrow_d 0$ 不对，因为 $X_n \pm Y_n$ 的方差为 $2\sigma^2$ ，矛盾

对于减法也一样

由上一页的定理最后一条可知，依分布收敛会对加减法不收敛，会给统计带来困扰，如：

$$t = \frac{\sqrt{n}(\bar{x} - \mu)}{s} = \frac{\xi_n}{\zeta_n}$$

或者对于统计量 $t_n = X_n + Y_n$ ，若 $X_n \rightarrow_d X, Y_n \rightarrow_d Y$ ，不能推出 $X_n + Y_n = t_n \rightarrow_d X + Y$

即显著性检验方面，不能确定其极限分布

则不能用极限分布去代替分布做假设检验

Theorem 2.1.2 (Slutsky's Theorem)

Let $X_n \rightarrow_d X, Y_n \rightarrow_p C$ (i.e. $Y_n \rightarrow_d C$). Then

1. $X_n + Y_n \rightarrow_d X + C$
2. $X_n Y_n \rightarrow_d CX$
3. $X_n / Y_n \rightarrow_d X/C$ if $C \neq 0$

意义：两个随机变量的极限分布，加强某个随机变量的收敛性，即加强 Y_n ，使得其依概率收敛到常数 C ，则可以推出部分依分布收敛的加减乘除封闭性

$$t = \frac{\sqrt{n}(\bar{x} - \mu)}{s} = \frac{\xi_n}{\zeta_n}$$

$$\xi_n = \sqrt{n}(\bar{x} - \mu) \rightarrow_d \sigma N(0, 1)$$

$$s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2 = \frac{1}{n-1} \sum Y_i \rightarrow_p \sigma^2 \quad \text{大数定律}$$

\downarrow

$$\zeta_n \rightarrow_p \sigma$$

$$\therefore \frac{\xi_n}{\zeta_n} \rightarrow_d \frac{\sigma N(0, 1)}{\sigma} = N(0, 1)$$

作业：证明定理2.1.2的第三条（从定义开始推导）

Let $X_n \rightarrow_d X$, $Y_n \rightarrow_p C$ (i.e. $Y_n \rightarrow_d C$)

$$Y_n \rightarrow_p C \quad \forall \epsilon > 0 \quad P\{|Y_n - C| > \epsilon\} \rightarrow 0 \quad n \rightarrow \infty$$

$$\forall \epsilon > 0, \forall \delta > 0, \exists N, \text{ while } n > N \\ p\{|Y_n - C| > \epsilon\} < \delta$$

$$X_n \rightarrow_d X \quad F_n(x) = P\{X_n \leq x\} \rightarrow P\{X \leq x\} = F(x) \quad n \rightarrow \infty$$

$$\forall \epsilon > 0, \forall \delta > 0, \exists N, \text{ while } n > N \\ |F_n(x) - F(x)| < \delta$$

3. $X_n/Y_n \rightarrow_d X/C$ if $C \neq 0$

为了证明 $X_n/Y_n \rightarrow_d X_n/C$, 首先证明 $1/Y_n \rightarrow_p 1/C$

$$\begin{aligned} & \forall \epsilon > 0 \\ P\left\{\left|\frac{1}{Y_n} - \frac{1}{C}\right| \geq \epsilon\right\} &= P\left\{\left|\frac{Y_n - C}{Y_n C}\right| \geq \epsilon\right\} \\ &= P\left\{\left|\frac{Y_n - C}{C^2 + C(Y_n - C)}\right| \geq \epsilon, |Y_n - C| < \epsilon\right\} \\ &\quad + P\left\{\left|\frac{Y_n - C}{C^2 + C(Y_n - C)}\right| \geq \epsilon, |Y_n - C| \geq \epsilon\right\} \\ &\leq P\left\{\left|\frac{Y_n - C}{C^2 - \epsilon|C|}\right| \geq \epsilon\right\} + P\{|Y_n - C| \geq \epsilon\} \\ &= P\{|Y_n - C| \geq (C^2 - \epsilon|C|)\epsilon\} + P\{|Y_n - C| \geq \epsilon\} \\ &\rightarrow 0 \quad (n \rightarrow \infty) \end{aligned}$$

即 $1/Y_n \rightarrow_p 1/C$

再利用定理第二条即可得 $X_n/Y_n \rightarrow_d X_n/C$

这个题也可以用连续影射定理来证明

Theorem 2.1.3 (Continuous Mapping Theorem) 连续映射定理, 若 定理2.1.2不满足, 则验证这个定理

Let X_1, X_2, \dots and X be k -dim random vectors, $g: \mathcal{R}^k \rightarrow \mathcal{R}$ be continuous. Then

1. $X_n \rightarrow_{a.s.} X \Rightarrow g(X_n) \rightarrow_{a.s.} g(X)$
2. $X_n \rightarrow_p X \Rightarrow g(X_n) \rightarrow_p g(X)$
3. $X_n \rightarrow_d X \Rightarrow g(X_n) \rightarrow_d g(X)$

例如:

$$\begin{matrix} \xi_n \rightarrow_d \xi \\ \zeta_n \rightarrow_d \zeta \end{matrix} \Rightarrow t_n = \frac{\xi_n}{\zeta_n} \rightarrow_d \frac{\xi}{\zeta}$$

用连续映射 $g(x, y) = \frac{x}{y} \quad y \neq 0$

$$t_n = g(\xi_n, \zeta_n) = \frac{\xi_n}{\zeta_n}$$

若能证明 $\begin{bmatrix} \xi_n \\ \zeta_n \end{bmatrix} \rightarrow_d \begin{bmatrix} \xi \\ \zeta \end{bmatrix}$

, 证明这个比较麻烦, 可以用两种方法:

- 方法一: 由定义, 联合分布的连续点收敛到点
- 方法二: 特征函数 (随机向量的特征函数 $\varphi(t) = Ee^{it'X_n}$)

首先验证定理2.1.2, 不满足再验证 定理2.1.3

Examples:

1. If $X_n \rightarrow_d N(0, 1)$, then $X_n^2 \rightarrow_d N^2(0, 1) = \chi_1^2$
2. If $(X_n, Y_n) \rightarrow_d N(0, 1)$, then $X_n/Y_n \rightarrow_d \text{Cauchy}$.

Proof.

Let $(X, Y) \rightarrow_d N(0, I)$, then X, Y are independent

$X/Y \rightarrow_d \text{Cauchy}$, and $P(g(X, Y)) := X/Y$ is discontinuous

Order relations

Order relations could greatly simplify writings in a proof.

Definition

Let $\{a_n\}$'s and $\{b_n\}$'s be two sequences.

- $a_n = O(b_n) \iff |a_n/b_n| \leq C \iff \limsup |a_n/b_n| < \infty$
- $a_n = o(b_n) \iff a_n/b_n \rightarrow 0$
- $a_n \sim b_n \iff a_n/b_n \rightarrow 1$

Definition [Stochastic orders]

Let $\{X_n\}$'s and $\{Y_n\}$'s be two sequences of r.v.'s

- $X_n = O_p(1)$
 $\iff \forall \epsilon > 0, \exists M_\epsilon, N_\epsilon, \forall n \geq N_\epsilon \quad P(|X_n| > M_\epsilon) < \epsilon$
 即超过 M_ϵ 的概率不超过 ϵ
 $\iff \forall \epsilon > 0, \exists M_\epsilon \quad \sup_n P(|X_n| > M_\epsilon) < \epsilon$
 $\iff \lim_{M \rightarrow \infty} \limsup_n P(|X_n| > M) = 0$ ("tightness")
- $X_n = O_p(Y_n)$ if $X_n/Y_n = O_p(1)$
- $X_n = o_p(1)$ if $X_n \rightarrow_p 0$
- $X_n = o_p(Y_n)$ if $X_n/Y_n = o_p(1)$ 分子收敛于0的速度比分母快

将数列的无穷小转化为概率的无穷小

在皮亚诺余项的泰勒公式里面有应用

Remark 2.2.1 If $X_n = O_p(1)$, we say that $\{X_n\}$ is "stochastically bounded", or is "tight". It means that no mass will escape to ∞ with positive probability.

随机有界：不存在能够逃到无穷大的点序列

随机不有界：...

例：不是 $O_p(1)$ 的序列 $\{x_n\}$

$$X_n = \begin{cases} n & \frac{1}{2} \\ 0 & \frac{1}{2} \end{cases}$$

令 $\epsilon = 0.1$

$$\begin{cases} p\{|X_n| > M_{0.1}\} < 0.1 \text{ 不成立} & \text{找 } M_{0.1} \\ \forall n > N \text{ 令 } N = 10000, n = 100000, p\{|X_{100000}| > M_{0.1}\} = \frac{1}{2} \text{ 不成立} & \text{找 } N \end{cases}$$

即有一个点概率会逃到无穷大

Theorem 2.2.1

- If $X_n \rightarrow_d X$, then $X_n = O_p(1)$
- If $X_n = o_p(1)$, then $X_n = O_p(1)$

Proof. 没看懂...似乎是数学分析的一个课后习题

例:

定义 $X_n = \frac{\sqrt{n}(\bar{x} - \mu)}{\sigma} \rightarrow_d N(0, 1)$

则 $X_n = O_p(1)$ 不会有概率逃到 ∞

则 $\frac{\bar{x} - \mu}{\sigma} = O_p\left(\frac{1}{\sqrt{n}}\right)$

由定义 $X_n = O_p(Y_n) \Rightarrow X_n/Y_n = O_p(1)$

则 $\bar{x} - \mu = O_p\left(\frac{1}{\sqrt{n}}\right)$ σ 是常数

大数定理: 样本均值很靠近总体均值

中心极限定理: 样本均值与总体均值的距离大致为 $\frac{1}{\sqrt{n}} \rightarrow$ 阶数

中心极限定理可以告诉我们更多的信息, 如果样本n为100, 则距离大致为0.1

Exercises 2.3

1. Check whether the following statements are true or false. With or prove your claim or given a counter-example.

a. If $X_n \rightarrow_p X$ and $Y_n \rightarrow_p Y \Rightarrow X_n Y_n \rightarrow_p XY$

b. If $X_n \rightarrow_{a.s.} X$ and $Y_n \rightarrow_{a.s.} Y \Rightarrow X_n Y_n \rightarrow_{a.s.} XY$

c. If $X_n \rightarrow X$ in L_r and $Y_n \rightarrow Y$ in $L_r (r > 0) \Rightarrow X_n Y_n \rightarrow XY$ in L_r

d. If $X_n \rightarrow_d X$ and $Y_n \rightarrow_d Y \Rightarrow X_n Y_n \rightarrow_d XY$

2. What happens if we replace the operation of multiplication by that of division in the last question?

3. Prove the equivalence of several definitions for $X_n = O_p(1)$ given in the notes.

作业: 证明

- $O_p(1) \times o_p(1) = o_p(1)$
- $O_p(1) \times O_p(1) = O_p(1)$

从定义出发

令

$$\begin{aligned}
X_n = o_p(1) &\Rightarrow X_n \rightarrow_p 0 \\
&\Rightarrow p\{|X_n| > \delta\} \rightarrow 0 \\
&\Rightarrow P\{|X_n| > \delta\} < \epsilon \quad \forall n \geq N_\epsilon \\
Y_n = O_p(1) &\Rightarrow \forall \epsilon > 0 \quad \exists M_\epsilon, N_\epsilon \\
&\quad P\{|Y_n| > M_\epsilon\} < \epsilon \\
Z_n = O_p(1) &\Rightarrow \forall \epsilon > 0 \quad \exists M_\epsilon, N_\epsilon \\
&\quad P\{|Z_n| > M_\epsilon\} < \epsilon
\end{aligned}$$

则 $\forall \delta > 0, \exists M_\epsilon, M_{\epsilon_1}, s. t.$

$$\begin{aligned}
P\{|X_n Y_n| > \delta\} &= P\{|X_n| \times |Y_n| > \delta\} \\
&= P\{|X_n| \times |Y_n| > \delta, |Y_n| > M_\epsilon\} + P\{|X_n| \times |Y_n| > \delta, |Y_n| \leq M_\epsilon\} \\
&\leq P\{M_\epsilon \times |X_n| > \delta\} + P\{|Y_n| > M_\epsilon\} \\
&\leq 2\epsilon
\end{aligned}$$

$$\begin{aligned}
P\{|Z_n Y_n| > M_\epsilon\} &= P\{|Z_n| \times |Y_n| > M_\epsilon\} \\
&= P\{|Z_n| \times |Y_n| > M_\epsilon, |Z_n| > M_{\epsilon_1}\} + P\{|Z_n| \times |Y_n| > M_\epsilon, |Z_n| \leq M_{\epsilon_1}\} \\
&\leq P\{|Z_n| > M_{\epsilon_1}\} + P\{M_{\epsilon_1} |Y_n| > M_\epsilon\} \\
&< 2\epsilon
\end{aligned}$$

Chapter 3

Central Limit Theorems and Related Refinements

Limit theorems for sums of independent r.v.'s occupy a special place in probability and mathematical statistics.

First of all, they are the simplest case to study.

Second, the theory for sums of independent r.v.'s are the most complete.

Third, studies on sum of independent r.v.'s can shed light on other classes of statistics such as the function of the sum of independent r.v.'s, U -statistics, L -statistics and symmetric statistics, etc.

中心极限定理在概率论和数理统计中由重要的地位，它简单、完整，而且还能引领其余统计学领域的发展，比如独立随机变量的和、 U 统计量、 L 统计量、对称统计量等等

In this chapter, we shall discuss some important issues concerning the sum of independent r.v.'s.

In particular, we shall discuss the asymptotic normality, Berry-Esseen bounds (uniform or nonuniform), Edgeworth expansions (uniform or nonuniform), large deviations and saddlepoint approximations, and others.

This will pave the way for other classes of statistics.

本章将讨论有关独立随机变量的总和的一些重要问题。

特别是，我们将讨论渐近正态性，Berry-Esseen边界（一致或不一致），Edgeworth展开（一致或不一致），大偏差和鞍点近似等。

这是其他类别的统计的基础。

3.1.1 CLT for i.i.d. r.v.s

独立随机变量的中心极限定理

When X_1, \dots, X_n are i.i.d., we get the following simple CLT.

Theorem 3.1.1 (levy theorem)

Let X_1, \dots, X_n be i.i.d. r.v.s with $EX_1 = 0, \sigma^2 = EX_1^2 < \infty$

Let $F_n(x) = P\left(\sqrt{n}\bar{X}/\sigma \leq x\right)$

Then

$$\sup_{x \in R} |F_n(x) - \Phi(x)| \rightarrow 0$$

$$\frac{\sqrt{n}\bar{x}}{\sigma} = \sum \left(\frac{x_i}{\sqrt{n}\sigma} \right) \triangleq \sum_{i=1}^n Y_i \quad \begin{cases} E(Y_i) = 0 \\ \text{Var}(Y_i) = \frac{1}{n} \\ \text{Var}(\sum Y_i) = 1 \end{cases}$$

3.1.2 CLT for triangular arrays with finite variances

The following theorem states that a sum of a large number of small independent effects has approximately a normal distribution.

大量小的独立效应之和大约具有正态分布。

Theorem 3.1.2 (Lindeberg-Feller CLT)

For each n ,

Let $X_{n,k}, 1 \leq k \leq n$ be independent r.v.s with $EX_{n,k} = 0$ and $\sum_{k=1}^n \sigma_{n,k}^2 := \sum_{k=1}^n EX_{n,k}^2 = 1$.

Denote $F_n(x) = P(\sum_{m=1}^n X_{n,k} \leq x)$.

Then the following two statements are equivalent.

1. The Lindeberg condition holds: 截尾二阶矩趋于0

$$\forall \epsilon > 0 : \lim_{n \rightarrow \infty} \sum_{k=1}^n EX_{n,k}^2 I\{|X_{n,k}| \geq \epsilon\} = 0$$

2. a. $\max_{1 \leq k \leq n} \sigma_{n,k}^2 \rightarrow 0$

b. $\sup_{x \in R} |F_n(x) - \Phi(x)| \rightarrow 0$

对于 $\sum_{k=1}^n \sigma_{n,k}^2 = 1$ 这个条件来说, 会存在两种情况

1. 所有 $\sigma_{n,k}^2$ 都差不多

2. 某一个或几个 $\sigma_{n,k}^2$ 相对于其余的来说比较大

而 $\max_{1 \leq k \leq n} \sigma_{n,k}^2 \rightarrow 0$ 这个就说明, 不会存在第二种情况。

Remark 3.1.1 Clearly, the Lindeberg condition implies that

$$\forall \epsilon > 0 : \lim_{n \rightarrow \infty} \sum_{m=1}^n P(|X_{n,k}| \geq \epsilon) = 0$$

which further implies that

$$\forall \epsilon > 0 : \sup_n P(|X_{n,k}| \geq \epsilon) \rightarrow 0$$

That is, all the individual terms $X_{n,k}$ are uniformly small.

三角列:

$$\begin{array}{ccccccc} x_{1,1} & & & & & & \rightarrow \sum_{k=1}^1 x_{n,k} \\ x_{2,1} & x_{2,2} & & & & & \rightarrow \sum_{k=1}^2 x_{n,k} \\ x_{3,1} & x_{3,2} & x_{3,3} & & & & \rightarrow \sum_{k=1}^3 x_{n,k} \\ \vdots & & & & & & \\ x_{n,1} & x_{n,2} & x_{n,3} & \cdots & x_{n,n} & \rightarrow \sum_{k=1}^n x_{n,k} \end{array}$$

在这个定理中, 只要求 $x_{n,k}$ 独立 (n 固定), 可以不同分布

期望为0, 方差为1不是一个限制条件, 可以自己构造一下

三角列的中心极限定理: 研究三角阵随机变量的行和的分布

相对于 Levy CLT, Lindeberg-Feller CLT 的好处是

1. 对行和不要求同分布
2. 给了一个充要条件

Theorem 3.1.3 Lyapunov CLT

For each n

Let $X_{n,k}, 1 \leq k \leq n$, be independent r.v.s with $EX_{n,k} = 0$ and $\sum_{k=1}^n \sigma_{n,k}^2 := \sum_{k=1}^n EX_{n,k}^2 = 1$.

Denote $F_n(x) = P(\sum_{m=1}^n X_{n,k} \leq x)$

If

$$\sum_{k=1}^n E|X_{n,k}|^{2+\delta} \rightarrow 0$$

Then

$$\sup_{x \in \mathbb{R}} |F_n(x) - \Phi(x)| \rightarrow 0$$

Proof.

The Lindeberg condition holds since

$$\begin{aligned}\sum_{k=1}^n EX_{n,k}^2 I\{|X_{n,k}| \geq \epsilon\} &\leq \frac{1}{\epsilon^\delta} \sum_{k=1}^n E|X_{n,k}|^{2+\delta} I\{|X_{n,k}| \geq \epsilon\} \\ &\leq \frac{1}{\epsilon^\delta} \sum_{k=1}^n E|X_{n,k}|^{2+\delta} \\ &\rightarrow 0\end{aligned}$$

Lyapunov 条件 $\sum_{k=1}^n E|X_{n,k}|^{2+\delta} \rightarrow 0$

代价: $2 + \delta$ 阶矩存在

而 L-F条件只需要2阶矩存在

省略一点东西...

An application: CLT for independent r.v.s

Lindeberg-Feller CLT

Lyapunov CLT

3.2 Uniform Berry-Esseen Bounds

给出了CLT下的收敛速度

X_1, X_2, \dots, X_n independent r.v.'s

$$E(X_j) = 0$$

$$E|X_j|^{2+\delta} < \infty \text{ for some } 0 < \delta \leq 1 (j = 1, 2, \dots)$$

$$EX_j^2 = \sigma_j^2 \quad B_n^2 = \sum_{j=1}^n \sigma_j^2 \quad L_{n,\delta} = \frac{\sum_{j=1}^n E|X_j|^{2+\delta}}{B_n^{2+\delta}}$$

In particular, in i.i.d. case, where $EX_j^2 = \sigma^2$, we have (独立同分布, 把上面的替换一下)

$$B_n^2 = n\sigma^2 \quad L_{n,\delta} = \frac{nE|X_1|^{2+\delta}}{(n\sigma^2)^{\frac{2+\delta}{2}}} = \frac{C}{n^{\delta/2}} \quad L_{n,1} = \frac{C}{n^{1/2}}$$

Theorem 3.2.1 (Berry-Essen bounds for independent r.v.'s) 独立的情况

Let X_1, X_2, \dots, X_n be independent r.v.'s such that $EX_j = 0$ and $E|X_j|^{2+\delta} < \infty$ for some $0 < \delta \leq 1$.

Put

$$L_{n,\delta} = B_n^{-(2+\delta)} \sum_{j=1}^n E|X_j|^{2+\delta} = \frac{\sum_{j=1}^n E|X_j|^{2+\delta}}{B_n^{2+\delta}}$$

Then for all n

$$\sup_{x \in \mathbb{R}} |F_n(x) - \Phi(x)| \leq AL_{n,\delta}$$

Theorem 3.2.2 (Berry-Essen bounds for iid. r.v.'s) 独立同分布的情况

Let X_1, \dots, X_n be i.i.d r.v.'s

$$\text{Let } \delta \in (0, 1] \quad EX_1 = 0 \quad EX_1^2 = \sigma^2 > 0 \quad E|X_1|^{2+\delta} < \infty \quad \rho_\delta = E|X_1|^{2+\delta} / \sigma^{2+\delta}$$

Then for all n ,

$$\sup_{x \in \mathbb{R}} |F_n(x) - \Phi(x)| \leq \frac{A\rho_\delta}{n^{\delta/2}}$$

由上面的情况可以立即推出引理3.2.2

Corollary 3.2.2 Let X_1, X_2, \dots, X_n be i.i.d. r.v.'s

$$\text{Let } \delta = 1 \quad EX_1 = 0 \quad EX_1^2 = \sigma^2 > 0 \quad E|X_1|^3 < \infty \quad \rho = E|X_1|^3 / \sigma^3$$

Then for all n ,

$$\sup_{x \in \mathbb{R}} |F_n(x) - \Phi(x)| \leq \frac{A\rho_\delta}{\sqrt{n}}$$

3.3 Non-Uniform Berry-Essen Bounds