Theorem 7.3.4 Law of Large Numbers

Assume event E occurs with probability x and let m denote the number of times E occurs in n trials. Let $\varepsilon>0$ and $\delta>0$ be given. The probability that $\frac{m}{n}$ differs from x by less than δ is greater than $1-\varepsilon$, ie.,

$$\mathcal{P}\left(\left|\tfrac{m}{n}-x\right|<\delta\right)>1-\varepsilon,$$

for all n sufficiently large.

Note: This does not say that E occurs exactly xn times, nor that E must occur roughly xn times.

Proof of Theorem 7.3.4

In terms of binomial polynomials, we want to show that given $\varepsilon, \delta>0$,

$$(7.3.7) \qquad \sum_{\substack{0\le m\le n \\ \left|\frac{m}{n}-x\right|<\delta}} P_{n,m}(x) > 1-\varepsilon \qquad \text{all } n \text{ large}$$

Since lower bounds are difficult in general we consider the complementary sum giving the probability of what we don't want

$$\sum_{\substack{0\le m\le n \\ |\frac{m}{n}-x|\ge\delta}} p_{n,m}(x) = 1 - \sum_{\substack{0\le m\le n \\ |\frac{m}{n}-x|<\delta}} p_{n,m}(x),$$

that we estimate as

$$\sum_{\substack{0\le m\le n \\ |\frac{m}{n}-x|\ge\delta}} p_{n,m}(x) \le \frac{1}{\delta^2} \sum_{\substack{0\le m\le n \\ |\frac{m}{n}-x|\ge\delta}} \left(\frac{m}{n}-x\right)^2 p_{n,m}(x)$$

$$\le \frac{1}{n^2\delta^2} \sum_{m=0}^{n} (m-nx)^2 p_{n,m}(x)$$

$$\le \frac{1}{n^2\delta^2}\left(\sum_{m=0}^{n} m^2 p_{n,m}(x) - 2nx\sum_{m=0}^{n} m\, p_{n,m}(x) + n^2x^2 \sum_{m=0}^{n} p_{n,m}(x)\right)$$

Using (7.3.4) - (7.3.6), the sums on the right simplify to $nx(1-x)$. Since $x(1-x)\le\frac{1}{4}$ for $0\le x\le 1$,

$$(7.3.8) \quad \sum_{\substack{0\le m\le n \\ |\frac{m}{n}-x|\ge\delta}} p_{n,m}(x) \le \frac{1}{4n\delta^2}$$

and

$$\sum_{\substack{0\le m\le n \\ |\frac{m}{n}-x|<\delta}} p_{n,m}(x) \ge 1 - \frac{1}{4n\delta^2}.$$

For given $\varepsilon, \delta > 0$, we can insure $(4n\delta^2)^{-1} < \varepsilon$

by choosing $n > \frac{1}{4\delta^2\varepsilon}$.

<u>Proof of Theorem 7.3.1</u>

We first define the approximating polynomial, named after the person who made this proof.

<u>Definition 7.3.3</u>

We partition $[0,1]$ by a uniform mesh with $n+1$ nodes $x_m = \frac{m}{n}$, $m = 0, 1, \ldots, n$.

The <u>Bernstein polynomial</u> of order n for f on $[0,1]$ is

$$B_n(f,x) = B_n(x) = \sum_{m=0}^{n} f(x_m)\, P_{n,m}(x).$$

Note that $\deg(B_n) \leq n$.

The reason that B_n approximates f is intuitive

$$B_n(x) = \sum_{x_m \approx x} f(x_m) P_{n,m}(x) + \sum_{\substack{|x_m - x| \\ \text{large}}} f(x_m)\, P_{n,m}(x),$$

The first sum converges to f as n increases because we can find $\frac{m}{n}$ arbitrarily close to x while the second sum goes to zero by the

Law of Large Numbers.

Example 7.3.5

Consider x^2 on $[0,1]$ with $n \geq 2$,

$$B_n(x) = \sum_{m=0}^{n} \left(\frac{m}{n}\right)^2 P_{n,m}(x).$$

Using the identities (7.3.4)-(7.3.6),

$$B_n(x) = \left(1-\frac{1}{n}\right)x^2 + \frac{1}{n}x.$$

Note that $B_n(x^2, x) \neq x^2$ and the error is

$$|x^2 - B_n(x)| = \frac{1}{n}x(1-x)$$

which tends to zero like $\frac{1}{n}$ on $[0,1]$. This contrasts with interpolating polynomials and Taylor polynomials, which both have the property that if f is a polynomial then $P_n = f$ for $n \geq \deg(f)$.

Example 7.3.6 For e^x on $[0,1]$,

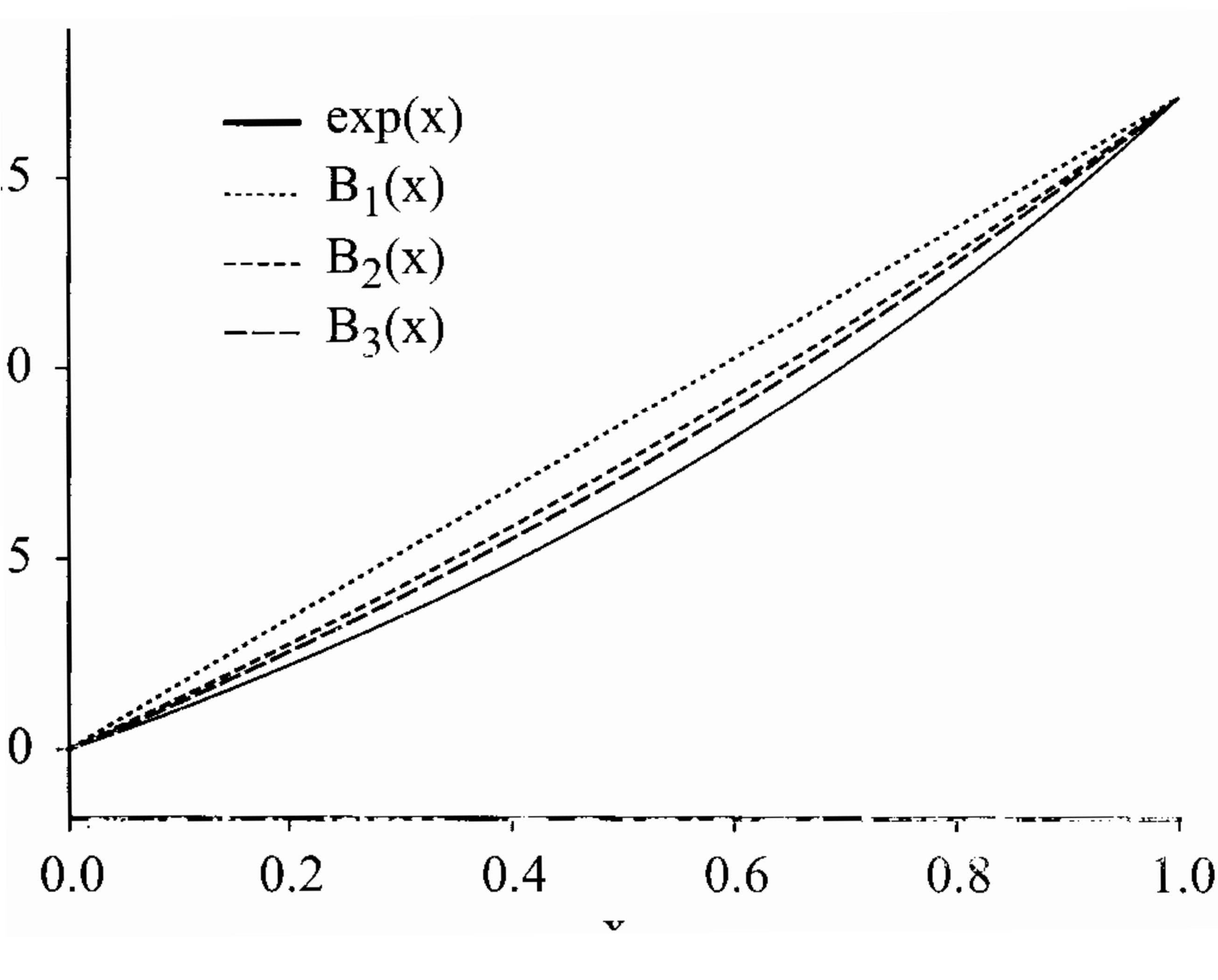
$$B_1(x) = (1-x) + ex$$

$$B_2(x) = (1-x)^2 + 2e^{1/2}x(1-x) + ex^2$$

$$B_3(x) = \dots$$

—plot—

We prove that given $\varepsilon > 0$, there is an n



exp(x)
B1(x)
B2(x)
B3(x)
.5
0
5
0
0.0
0.2
0.4
0.6
0.8
1.0

such that

$$\sup_{0 \le x \le 1} |f(x) - B_n(x)| < \varepsilon.$$

Using (7.3.4), we write

$$f(x) - B_n(x) = \sum_{m=0}^{n} f(x)\, p_{n,m}(x) - \sum_{m=0}^{n} f(x_m)\, p_{n,m}(x)$$

$$= \sum_{m=0}^{n} (f(x) - f(x_m))\, p_{n,m}(x).$$

We expect that we can make $f(x) - f(x_m)$ small when x is close to x_m by continuity. For $\delta > 0$, we write

(7.3.9)

$$f(x) - B_n(x) = \sum_{\substack{0 \le m \le n \\ |x - x_m| < \delta}} (f(x) - f(x_m))\, p_{n,m}(x)$$

$$+ \sum_{\substack{0 \le m \le n \\ |x - x_m| \ge \delta}} (f(x) - f(x_m))\, p_{n,m}(x).$$

Theorem 6.4.2 implies f is uniformly continuous on $[0,1]$. Given $\varepsilon > 0$, there is a $\delta > 0$ such that

$$|f(x) - f(x_m)| < \varepsilon/2$$

for all x, x_m in $[0,1]$ with $|x - x_m| \le \delta$. Given δ, by the way, we can find x_m such that $|x - x_m| \le \delta$

for all sufficiently large n, since the rationals are dense in $[0,1]$.

Thus,

$$\left| \sum_{\substack{0 \le m \le n \\ |x-x_m| < \delta}} (f(x) - f(x_m)) \, p_{n,m}(x) \right|$$

$$\le \sum_{\substack{0 \le m \le n \\ |x-x_m| < \delta}} |f(x) - f(x_m)| \; p_{n,m}(x)$$

$$< \frac{\varepsilon}{2} \sum_{0 \le m \le n} p_{n,m}(x) = \frac{\varepsilon}{2}.$$

Now the second sum on the right in (7.3.9) is bounded after we realize that Theorem 6.5.4 implies $|f|$ is bounded on $[0,1]$ by some constant M. Hence, (7.3.8) implies

$$\left| \sum_{\substack{0 \le m \le n \\ |x-x_m| \ge \delta}} (f(x) - f(x_m)) \, p_{n,m}(x) \right| \le 2M \sum_{\substack{0 \le m \le n \\ |x-x_m| \ge \delta}} p_{n,m}(x)$$

$$\le \frac{M}{2n\delta^2}.$$

Given δ from the first estimate, we can force $\frac{M}{2n\delta^2} < \frac{\varepsilon}{2}$ by taking n sufficiently large.

Book: how general should §7.4 be presented – or transition better

§7.4 Compact Sets in $C([a,b])$.

Example 7.2.1 shows that $C([a,b])$ is not compact.

Example 7.4.1

The sequence $\{x^n, n \in \mathbb{N}\}$ does not have a convergent subsequence in $C([a,b])$, i.e., there is no subsequence that converges uniformly.

– start –

This raises the issue of describing the compact subsets of $C([a,b])$.

By Theorem 3.2.1, if $K \subset C([a,b])$ is compact, then K is closed and bounded. Closed is rather obvious: K is closed if for any sequence of functions $\{f_n\}$ in K that converges to f in the metric of $C([a,b])$ (uniformly), we have $f \in K$.

Example 7.4.2 Let $F = \{f \in C([a,b]) \mid \sup_{a \le x \le b} |f(x)| < 1\}$.