Bernstein’s Theorem

Binomial identities. Verify that, for every real \( x \),
\[
\sum_{k=0}^{n} \binom{n}{k} x^k (1 - x)^{n-k} = 1,
\]
\[
\sum_{k=0}^{n} k \binom{n}{k} x^k (1 - x)^{n-k} = nx,
\]
\[
\sum_{k=0}^{n} (k - nx)^2 \binom{n}{k} x^k (1 - x)^{n-k} = nx(1 - x).
\]
These are the total probability, expected value, and variance for the binomial distribution (\( n \) independent trials with probability of success \( x \in [0, 1] \)).

Theorem [Weierstrass, 1885]. Let \( f \) be continuous on \([a, b]\). Then for every \( \varepsilon > 0 \) there is a polynomial \( p \) such that \(|f(x) - p(x)| < \varepsilon \), \( a \leq x \leq b \).

Thus every continuous function on a compact interval is a uniform limit of polynomials.

The approach of Weierstrass was to establish
\[
\frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} f(s) e^{-\left(\frac{s}{\delta}\right)^2} ds \to f(t) \quad (\delta \to 0^+),
\]
so that the partial sums of the power series for the convolution integral approximate \( f \).

Theorem [Bernstein, 1912]. If \( f \) is continuous on \([0, 1]\), the polynomials
\[
B_n(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1 - x)^{n-k}
\]
converge to \( f \) uniformly on \([0, 1]\).

Proof. Given \( \varepsilon > 0 \), choose \( \delta > 0 \) so that \(|f(x) - f(x')| < \varepsilon/2\) whenever \(|x - x'| < \delta\).

Note that
\[
\sum_{|\frac{k}{n} - x| \geq \delta} \binom{n}{k} x^k (1 - x)^{n-k} \leq \frac{1}{\delta^2 \sqrt{\pi}} \sum_{|\frac{k}{n} - x| \geq \delta} (\frac{k}{n} - x)^2 \binom{n}{k} x^k (1 - x)^{n-k} = \frac{x(1-x)}{n\delta^2}.
\]
Consequently, if \( M \) is an upper bound for \(|f|\),
\[
|f(x) - B_n(x)| = \left| \sum_{k=0}^{n} (f(x) - f\left(\frac{k}{n}\right)) \binom{n}{k} x^k (1 - x)^{n-k} \right|
\leq \sum_{|\frac{k}{n} - x| < \delta} |f(x) - f\left(\frac{k}{n}\right)| \binom{n}{k} x^k (1 - x)^{n-k} + \sum_{|\frac{k}{n} - x| \geq \delta} |f(x) - f\left(\frac{k}{n}\right)| \binom{n}{k} x^k (1 - x)^{n-k}
\leq \frac{\varepsilon}{2} \sum_{k=0}^{n} \binom{n}{k} x^k (1 - x)^{n-k} + 2M \frac{x(1-x)}{n\delta^2} \leq \frac{\varepsilon}{2} + \frac{M}{2n\delta^2}.
\]
It remains to choose \( n > M/\varepsilon \delta^2 \). \( \square \)