# Large numbers for statistics

Sam Cohen
University of Oxford

December 5, 2019

**Abstract**

Lot's of statistical mehtods are based on the law of large numbers so it makes sense to try and understand it a bit better. We will see a proof of the basic law of large numbers and the central limit theorem, and then some applications to hypothesis testing.

## 1 Introduction

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and $L^p$ the associated $L^p$ spaces. We want to understand the behaviour of $\bar{X}_n$ where $\bar{X}_n = \frac{1}{n}\sum X_i$. As $X_i$ is iid, we can do this by using the law of large numbers, which comes from Chebysheff's inequality.

**Theorem 1.1** (Law of Large Numbers). *As $n \to \infty$ we have $X_n \to \mu$.*

*Proof.* We seek to prove the weak law of large numbers, for a sequence of square-integrable independent random variables with common mean and common bound on the variance. To do this we use the simple calculations

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

and $V(\bar{X}_n) = \frac{1}{n^2}\sum V(X_i) \leq \bar{\sigma}^2/n$. This means that we can see

$$P(|\bar{X}_n - \mu| > \epsilon) \leq \bar{\sigma}^2/(\epsilon n) \to 0$$

as $n \to \infty$ for each $\epsilon$, is Tchebyshev's inequality. □

The use of laws of Large numbers dates back to the work of Jacob Bernoulli, but a key advance is the use of Stirling's formula to approximate the behaviour of the relevant integral by de Moivre, who proved the celebrated central limit theorem.

**Lemma 1.** *If two probability measures $\mathbb{P}$, $\mathbb{Q}$ supported on $\mathbb{R}$ have the same characteristic function, then they agree. If the characteristic function of one approaches that of the other then the measures must also be converging weakly.*

*Proof.* For each of $\mathbb{P}$ and $\mathbb{Q}$ we can find a cdf $F$,$G$. Then the characteristic function is the Fourier transform of the cdf, and the invertibiltiy of the transform shows the result. □

**Theorem 1.2** (Central limit theorem)**.** *$\xi_n = (\bar{X}_n - \mu)/(\sigma\sqrt{n})$. Then $\xi_n$ converges in distribution to $N(0,1)$.*

*Proof.* We compute the characteristic function using the

$$e^x = 1 + x + x^2/2 + x^3/3! + \frac{x^4}{24} + ...$$

$$(1 + t/n)^n \to e^t$$

$$E[e^{it\xi_n}] = (E[e^{itY_i/n}])^n = E[(1+itY_i-t^2Y_i^2+o(n^{-2})]^n = (1-t^2+o(n^{-2})^n = \exp(1-t\xi_n^2+o(n^{-2}))$$

Ignoring the $o(n^{-2})$ term, we get the characteristic function of a normal distribution, so by our lemma we have the normal limit. In the above $Y_i = (X_i - \mu)/\sqrt{n}$. □

This is closely related to Lévy's characterization of Brownian motion.