<div align="center">

Appendix
Class 6, 18.05, Spring 2014
Jeremy Orloff and Jonathan Bloom

</div>

1 Introduction

In this appendix we give more formal mathematical material that is not strictly a part of 18.05. This will not be on homework or tests. We give this material to emphasize that in doing mathematics we should be careful to specify our hypotheses completely and give clear deductive arguments to prove our claims. We hope you find it interesting and illuminating.

2 Proof the histogram will probably match the pdf:

We stated that one consequence of the law of large numbers is that as the number of samples increases the histogram of the samples has an increasing probability of matching the graph of the underlying pdf or pmf. We prove that here by applying the LoLN to each bin in the histogram.

Define an indicator (Bernoulli) random variable B_k for each bin, which is 1 if X is in bin k and 0 otherwise. The mean of B_k is then the probability that X is in bin k. So by the law of large numbers, if we take enough samples of X (and therefore of each B_k), then with high probability, the proportion of the samples in bin k will be arbitrarily close to the area of $f(x)$ over bin k.

3 The Chebyshev inequality

One proof of the LoLN follows from the following key inequality.

The Chebyshev inequality. Suppose Y is a random variable with mean μ and variance σ^2. Then for any positive value a, we have

$$P(|Y - \mu| \geq a) \ \leq \ \frac{\text{Var}(Y)}{a^2}.$$

In words, the Chebyshev inequality says that the probability that Y differs from the mean by more than a is bounded by $\text{Var}(Y)/a^2$. Morally, the smaller the variance of Y, the smaller the probability that Y is far from its mean.

Proof of the LoLN: Since $\text{Var}(\bar{X}_n) = \text{Var}(X)/n$, the variance of the average $\bar{X}_n$ goes to zero as n goes to infinity. So the Chebyshev inequality for $Y = \bar{X}_n$ and fixed a implies that as n grows, the probability that $\bar{X}_n$ is farther than a from μ goes to 0. Hence the probability that $\bar{X}_n$ is within a of μ goes to 1, which is the LoLN.

Proof of the Chebyshev inequality: The proof is essentially the same for discrete and continuous Y. We'll assume Y is continuous and also that $\mu = 0$, since replacing Y by

$Y - \mu$ does not change the variance. So

$$P(|Y| \geq a) = \int_{-\infty}^{-a} f(y)\,dy + \int_{a}^{\infty} f(y)\,dy \leq \int_{-\infty}^{-a} \frac{y^2}{a^2} f(y)\,dy + \int_{a}^{\infty} \frac{y^2}{a^2} f(y)\,dy$$

$$\leq \int_{-\infty}^{\infty} \frac{y^2}{a^2} f(y)\,dy = \frac{\mathrm{Var}(Y)}{a^2}.$$

The first inequality uses that $y^2/a^2 \geq 1$ on the intervals of integration. The second inequality follows because including the range $[-a, a]$ only makes the integral larger, as the integrand is positive.

18.05 Introduction to Probability and Statistics

Spring 2014