Please provide complete and well-written solutions to the following exercises.

Due November 16, in the discussion section.

Homework 6

Exercise 1. This exercise demonstrates that geometry in high dimensions is different than geometry in low dimensions.

Let $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$. Let $||x|| := \sqrt{x_1^2 + \cdots + x_n^2}$. Let $\varepsilon > 0$. Show that for all sufficiently large n, "most" of the cube $[-1, 1]^n$ is contained in the annulus

$$A := \{x \in \mathbf{R}^n : (1 - \varepsilon)\sqrt{n/3} \leq ||x|| \leq (1 + \varepsilon)\sqrt{n/3}\}.$$

That is, if $X_1, \ldots, X_n$ are each independent and identically distributed in $[-1, 1]$, then for n sufficiently large

$$\mathbf{P}((X_1, \ldots, X_n) \in A) \geq 1 - \varepsilon.$$

(Hint: apply the weak law of large numbers to $X_1^2, \ldots, X_n^2$.)

Exercise 2. Let $f, g, h : \mathbf{R} \to \mathbf{R}$. We use the notation $f(t) = o(g(t))\ \forall\ t \in \mathbf{R}$ to denote $\lim_{t\to 0} \left|\frac{f(t)}{g(t)}\right| = 0$. For example, if $f(t) = t^3\ \forall\ t \in \mathbf{R}$, then $f(t) = o(t^2)$, since $\lim_{t\to 0} |\frac{f(t)}{t^2}| = \lim_{t\to 0} |t| = 0$. Show: (i) if $f(t) = o(g(t))$ and if $h(t) = o(g(t))$, then $(f + h)(t) = o(g(t))$. (ii) If c is any nonzero constant, then $o(cg(t)) = o(g(t))$. (iii) $\lim_{t\to 0} g(t)o(1/g(t)) = 0$. (iv) $\lim_{t\to 0} o(g(t))/g(t) = 0$. (v) $o(g(t) + o(g(t))) = o(g(t))$.

Exercise 3 (Confidence Intervals)**.** Among 625 members of a bank chosen uniformly at random among all bank members, it was found that 25 had a savings account. Give an interval of the form $[a, b]$ where $0 \leq a, b \leq 625$ are integers, such that with about 95% certainty, the number of any set of 625 bank members with savings accounts chosen uniformly at random lies in the interval $[a, b]$. (Hint: if Y is a standard Gaussian random variable, then $\mathbf{P}(-2 \leq Y \leq 2) \approx .95$.)

Exercise 4 (Hypothesis Testing)**.** Suppose we run a casino, and we want to test whether or not a particular roulette wheel is biased. Let p be the probability that red results from one spin of the roulette wheel. Using statistical terminology, "$p = 18/38$" is the null hypothesis, and "$p \neq 18/38$" is the alternative hypothesis. (On a standard roulette wheel, 18 of the 38 spaces are red.) For any $i \geq 1$, let $X_i = 1$ if the i^{th} spin is red, and let $X_i = 0$ otherwise.

Let $\mu := \mathbf{E}X_1$ and let $\sigma := \sqrt{\mathrm{var}(X_1)}$. If the null hypothesis is true, and if Y is a standard Gaussian random variable

$$\lim_{n\to\infty} \mathbf{P}\left(\left|\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}\right| \geq 2\right) = \mathbf{P}(|Y| \geq 2) \approx .05.$$

To test the null hypothesis, we spin the wheel n times. In our test, we reject the null hypothesis if $|X_1 + \cdots + X_n - n\mu| > 2\sigma\sqrt{n}$. Rejecting the null hypothesis when it is true is called a type I error. In this test, we set the type I error percentage to be 5%. (The type I error percentage is closely related to the p-value.)

Suppose we spin the wheel $n = 3800$ times and we get red 1868 times. Is the wheel biased? That is, can we reject the null hypothesis with around 95% certainty?

Exercise 5. Suppose random variables $X_1, X_2, \ldots$ converge in probability to a random variable X. Prove that $X_1, X_2, \ldots$ converge in distribution to X.

Then, show that the converse is false.