Homework 1

MATH 166 - Fall 2024

Tufts University, Department of Mathematics

Due: September 12, 2024

1. Book Questions

Wasserman: Chapter 1: ~~#4~~, ~~#8~~; Chapter 2: ~~#14~~; Chapter 3: ~~#7~~; Chapter 4: #3

2. Supplemental Question (Experimentally Verifying the Law of Large Numbers)

The following may be performed in any scripting environment you prefer (MATLAB, R, Python, Julia,...)

(a) For $n = 10, 20, 30, \ldots, 10000$, sample n i.i.d. samples from $\mathcal{N}(0, 1)$ i.e. the random variable X with density $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$. Let $\bar{x}_n$ be the corresponding sample average. Plot $\bar{x}_n$ as a function of n. Describe the behavior as n increases. What does the Law of Large Numbers suggest will happen as $n \to \infty$?

(b) For $n = 10, 20, 30, \ldots, 10000$, sample n i.i.d. samples from the Cauchy distribution, i.e. the random variable X with density $f_X(x) = \frac{1}{\pi(1+x^2)}$. Let $\bar{x}_n$ be the corresponding sample average. Plot $\bar{x}_n$ as a function of n. Describe the behavior as n increases. What does the Law of Large Numbers suggest will happen as $n \to \infty$?

#4 1) show $\left(\bigcup_{i\in I} A_i\right)^c = \bigcap_{i\in I} A_i^c$

proof: let $I = \{1, 2 \ldots n\}$

$\left(\bigcup_{i\in I} A_i\right)^c = \{x \mid x \notin (A_1 \text{ or } A_2 \ldots \text{ or } A_n)\}$

$= \{x \mid x \notin A_1 \text{ or } x \notin A_2 \ldots \text{or } x \notin A_n\}$

theorem: de morgen

$= \{x \mid x \in A_1^c \text{ and } x \in A_2^c \text{ and } x \in A_n^c\} = \bigcap_{i\in I} A_i^c$

Therefore, $\left(\bigcup_{i\in I} A_i\right)^c = \bigcap_{i\in I} A_i^c$

since n is arbitrary, this will hold for any arbitrary index I ◻

2) show $\left(\bigcap_{i\in I} A_i\right)^c = \bigcup_{i\in I} A_i^c$

proof: let $I = \{1, 2, 3 \ldots n\}$

$\left(\bigcap_{i\in I} A_i\right)^c = \{x \mid x \notin (A_1 \cap A_2 \ldots \cap A_n)\}$

theorem: de morgen

$\Rightarrow \{x \mid x \in (A_1 \cup A_2 \ldots \cup A_n)^c\} = \bigcup_{i\in I} A_i^c$

therefore $\bigcup_{i\in I} A_i^c = \bigcup_{i\in I} A_i^c$ ◻

#8 $P\left(\bigcap_{i=1}^{\infty} A_i\right) = 1 - P\left(\left(\bigcap_{i=1}^{\infty} A_i\right)^c\right)$

$\geq 1 - \sum_{i}^{\infty} P(A_i^c) = 1 - 0 = 0$ $\square$

#2.14

CDF: $F(r) = P(\sqrt{x^2+y^2} \leq r)$

In other words, find chance that our distance from from origin is $\leq r$

total area of unit circle = π

total area of $r = \pi r^2$

$p(R \leq r) = \frac{\pi r^2}{\pi} = r^2$

PDF = $\frac{d}{dx}$ cdf $= 2r$

Q 3.7

given $P(X > 0) = 1$

show $E(x) = \int_0^\infty 1 - F(x)\,dx$

$$= \int_0^\infty 1\,dx - \int_0^\infty F(x)\,dx = x\Big|_0^\infty - \int_0^\infty F(x)\,dx$$

since $\int_0^\infty F\,dx$ $\quad u = F \quad du = f(x)dx \quad dv = dx \quad v = x$

$$= xF\Big|_0^\infty - \int_0^\infty x f(x)\,dx$$

putting it together, we get

$$\int_0^\infty 1 - F(x)\,dx = x\Big|_0^\infty - xF\Big|_0^\infty + \int_0^\infty x f(x)\,dx$$

$$= \underbrace{x(1 - F(x))\Big|_0^\infty}_{\text{goes to 0 (given)}} + \int_0^\infty x f(x)\,dx = E(x) \quad \square$$

4.3

$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$

since $X \sim$ bernoulli(p), that means that $0 \leq E(X) \leq 1 = p$

$var(x) = p(1-p)$, which is maximized when $p = 0.5$,

thus $var(X) \leq 0.25$ (bounded). Also, each trial is independent

since all conditions are satisfied, we can apply chebyshevs inequality:

$P(|\overline{X} - \mu| \geq \epsilon) \leq \frac{1}{\epsilon^2} var(x)$ where $\mu = p$ and $var(X) = \frac{p(1-p)}{n}$

$\rightarrow P(\overline{X} - \mu \geq \epsilon) \leq \frac{p(1-p)}{n\epsilon^2}$

hoeffding: $P(\overline{X} - \mu \geq \epsilon) \leq 2e^{(-2n\epsilon^2)} = \frac{2}{e^{2n\epsilon^2}}$

for large n, exponentials will decay faster than $\frac{1}{n}$

2a and 2b

for normal:

```
(base) benlam@Bens-MBP stats class % /opt/anaconda3/envs/nnfs/bin/python "/Users/benlam/Desktop/stats class/hw1 cauchy.py"
Average of 10 = -0.1845097282789873
Average of 20 = -0.42165020237398376
Average of 50 = 3.4129492291115806
Average of 100 = 1.8337337252330468
Average of 1000 = -1.3785455522043872
Average of 5000 = -0.03397670676338607
Average of 10000 = -1.881501663732055
```

For cauchy:

```
(base) benlam@Bens-MBP stats class % /opt/anaconda3/envs/nnfs/bin/python "/Users/benlam/Desktop/stats class/hw1.py"
Average of 10 = -0.3508669655471559
Average of 20 = 0.1674177802387504
Average of 30 = 0.08903820686967139
Average of 50 = -0.07005648099369158
Average of 10000 = 0.0013352045293414164
```

For both, the LoLN says that as $n \to \infty$, the average will converge to the expected value of the distribution