HOMEWORK 1

1) a) WTS: $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

↳ we can construct 3 disjoint events from the events A and B as follows...

$$\begin{cases} A^* = A \setminus B \\ B^* = B \setminus A \\ C = A \cap B \end{cases}$$

Then, we have $\mathbb{P}(A \cup B) = \mathbb{P}(A^* \cup B^* \cup C) = \mathbb{P}(A^*) + \mathbb{P}(B^*) + \mathbb{P}(C)$

$$= \mathbb{P}(A) - \mathbb{P}(A \cap B) + \mathbb{P}(B) \cancel{- \mathbb{P}(A \cap B)} \cancel{+ \mathbb{P}(A \cap B)}$$

$$= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \; \square.$$

b) H := passes house, S := passes senate. We have $\mathbb{P}(H) = 0.6$, $\mathbb{P}(S) = 0.8$, $\mathbb{P}(H \cup S) = 0.9$

↳ By additive property... $\mathbb{P}(H \cap S) = \mathbb{P}(H) + \mathbb{P}(S) - \mathbb{P}(H \cup S)$

$$= 0.6 + 0.8 - 0.9 = \boxed{0.5}$$

2) + := positive test, D := diseased; $\mathbb{P}(+ \mid D) = 0.9$, $\mathbb{P}(+ \mid D^c) = 0.15$. $\mathbb{P}(D \mid +) = ?$

↳ $$\mathbb{P}(D \mid +) = \frac{\mathbb{P}(+ \mid D) \cdot \mathbb{P}(D)}{\mathbb{P}(+)} = \frac{\mathbb{P}(+ \mid D) \cdot \mathbb{P}(D)}{\mathbb{P}(+ \mid D) \cdot \mathbb{P}(D) + \mathbb{P}(+ \mid D^c) \cdot \mathbb{P}(D^c)} = \frac{(0.9)(0.07)}{(0.9)(0.07) + (0.15)(0.93)}$$

$$= \frac{0.063}{0.063 + 0.1395} = \boxed{0.311}$$

3) a) Suppose $\mathbb{P}(\text{win}) = p$, and X is the number of games played to get 1st win. Then,
$X \sim \text{Geom}(p) \Rightarrow \mathbb{E}(X) = 1/p$. Assuming ad is true, we have $1/p = 3 \Rightarrow p = 1/3$

↳ Thus, $\mathbb{P}(\text{4 losses in 4 games}) = (1-p)^4 = (2/3)^4 = \boxed{16/81}$

b) If $\mathbb{E}(X) = 5$, then true probability $p^* = 1/5$. Then, $\mathbb{P}(\text{incorrectly accepting}) = \mathbb{P}(\text{winning w/in 1st 4 games})$

$$= \mathbb{P}(X \leq 4) = \sum_{k=1}^{4} (1-p)^{k-1} p$$

$$\frac{1}{5} \sum_{k=1}^{4} (4/5)^{k-1} = \frac{1}{5} \sum_{k=0}^{4} (4/5)^k = \frac{1}{5}\left(\frac{1-(4/5)^4}{1-4/5}\right) = \cancel{\tfrac{1}{5}}\left(\frac{1 - 256/625}{\cancel{1/5}}\right)$$

$$= 369/625 = \boxed{0.5904}$$

4) a) $$\mathbb{E}(X) = \sum_{k=-1}^{1} k \cdot \mathbb{P}(X = k) = \left(-1 \cdot \tfrac{1}{18}\right) + \cancel{\left(0 \cdot \tfrac{16}{18}\right)}^{0} + \left(1 \cdot \tfrac{1}{18}\right) = \boxed{0}$$

$$\text{Var}(X) = \mathbb{E}(X^2) - \cancel{\mathbb{E}(X)^2}^{0} = \mathbb{E}(X^2) = \left((-1)^2 \cdot \tfrac{1}{18}\right) + \cancel{\left(0^2 \cdot \tfrac{16}{18}\right)}^{0} + \left(1^2 \cdot \tfrac{1}{18}\right) = \boxed{1/9}$$

b) $\mathbb{P}(|X - \mathbb{E}(X)| \geq a) = \dfrac{\text{Var}(X)}{a^2} \Rightarrow \mathbb{P}(|X| \geq a) = \dfrac{1}{9a^2}$ so let $a = 1$ and...

$$\mathbb{P}(|X| \geq 1) \overset{\checkmark}{=} \frac{1}{9} \quad \text{thus} \quad \boxed{a = 1}$$

5) Markov's Inequality states that for any $a>0$, $\mathbb{P}(X \geq a) \leq \frac{E(X)}{a}$ where X is non-negative R.V.

a) Let $X :=$ # of H in 100 flips $\Rightarrow Y :=$ # T's in 100 flips. Note that $X+Y=100$, so... $\mathbb{P}(X \leq a) = \mathbb{P}(Y \geq 100-a)$

$\hookrightarrow \mathbb{P}(X \leq 10) = \mathbb{P}(Y \geq 90) \leq \frac{100(0.5)}{90} = 5/9$

so... $\boxed{\mathbb{P}(X \leq 10) \leq 5/9}$

b) Similar setup as part a, but $E(X) = 100(0.2) = 20$, $E(Y) = 100(0.8) = 80$. Still $\mathbb{P}(X \leq a) = \mathbb{P}(Y \geq 100-a)$.

$\hookrightarrow \mathbb{P}(X \leq 10) = \mathbb{P}(Y \geq 90) \leq \frac{80}{90} = 8/9$

so... $\boxed{\mathbb{P}(X \leq 10) \leq 8/9}$

Chebyshev's Inequality is an application of Markov's Inequality using variance. $\mathbb{P}(|X-\mu| \geq k) \leq \frac{\sigma^2}{k^2}$

c) Suppose $Y :=$ # T's in 100 tosses so $Y \sim \text{Binom}(100, 0.5) \Rightarrow E(Y) = 50$, $Var(Y) = np(1-p) = 25$

$\hookrightarrow \mathbb{P}(|Y-\mu| \geq k) \leq \frac{Var(Y)}{k^2}$. So if $40 \leq X \leq 60 \Rightarrow Y \leq 40$ or $Y \geq 60$ so $|Y-50| \geq 10$

$\Rightarrow \mathbb{P}(|Y-50| \geq 10) \leq \frac{25}{100} = 0.25$ so... $\boxed{\mathbb{P}(|X-50| \leq 10) \leq 0.25}$

d) Use same setup as part c, with $E(Y) = 80$, $Var(Y) = 16$

$\hookrightarrow \mathbb{P}(|Y-80| \geq 10) \leq \frac{16}{100} = 0.16$ so... $\boxed{\mathbb{P}(|X-20| \leq 10) \leq 0.16}$

6) a) Suppose f_X is the pdf of X. Then, $f_Y = f_X(\sqrt[3]{y^?}) \cdot \left|\frac{d}{dy}\sqrt[3]{y}\right| = f_X(y^{1/3}) \cdot \left|\frac{1}{3}y^{-2/3}\right|$

$\hookrightarrow f_X(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-x^2/2} \Rightarrow f_X(\sqrt[3]{y}) = \frac{1}{\sqrt{2\pi}} e^{-(y^{1/3})^2/2}$

$= \frac{e^{-y^{2/3}/2}}{\sqrt{2\pi}}$

Note that $D_X := (-\infty, \infty)$, $D_{y^{-1}} := (-\infty, \infty)$ so $D_{X \circ y^{-1}} := (-\infty, \infty)$. However, $\left|\frac{d}{dy} g^{-1}(y)\right|$ supported on $\mathbb{R}$ s.t. $y \neq 0$.

Thus, $\boxed{f_Y = \frac{1}{3\sqrt{2\pi}} \cdot e^{-\frac{1}{2}y^{2/3}} \cdot |y|^{-2/3}, \text{ supported on } y \in (-\infty, 0) \cup (0, \infty)}$

b) Split into 2 functions on 2 domains... $D_1 := [0, \infty)$ s.t. $Z = X$ for $X \geq 0$ and $D_2 := (-\infty, 0)$ s.t. $Z = -X$

$\hookrightarrow$ so only need to find f_Z for D_2. $f_Z = f_X(-z) \cdot \left|\frac{d}{dz}(-z)\right|$

$= f_X(z) \cdot |-1|$ since f_X is an even function.

$= f_X(z)$.

$\Rightarrow f_Z = f_X(z) + f_X(z) = 2f_X(z)$ for $z \in [0, \infty)$

Thus, $\boxed{f_Z = \frac{2}{\sqrt{2\pi}} \cdot e^{-z^2/2} \text{ supported on } z \in [0, \infty)}$

7) Let $\mathbb{1}_i = \begin{cases} 1 & \text{if elevator stops at floor } i \text{ w.p. } 1-p \\ 0 & \text{if elevator skips floor } i \text{ w.p. } p \end{cases}$

↳ Note that if $\mathbb{1}_i = 0$, then none of the 10 people chose that floor. Thus, $p = (9/10)^m$

$$\Rightarrow \mathbb{1}_i = \begin{cases} 1 & \text{w.p. } 1-(9/10)^m \\ 0 & \text{w.p. } (9/10)^m \end{cases}$$

Then, $\mathbb{E}\left(\sum_{i=1}^{10} \mathbb{1}_i\right) = \sum_{i=1}^{10} \mathbb{E}(\mathbb{1}_i) = \sum_{i=1}^{10} (1-(9/10)^m) = \boxed{10\left(1-(9/10)^m\right)}$

↑ by linearity of expectation

8) If A, B, C not mutually independent given conditions then $\mathbb{P}(A \cap B \cap C) \neq \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C)$

↳ Suppose we flip two independent fair coins. We will define RVs as follows...

$$X := \begin{cases} 1 & \text{if coin 1 heads} \\ 0 & \text{if coin 1 tails} \end{cases} \quad Y := \begin{cases} 1 & \text{if coin 2 heads} \\ 0 & \text{if coin 2 tails} \end{cases} \quad C := |A-B|$$

We can see that...

$$Z := \begin{cases} (1,0) \\ (0,1) \end{cases} > 1 \text{ w.p. } 1/2 \qquad \begin{cases} (1,1) \\ (0,0) \end{cases} > 0 \text{ w.p. } 1/2$$

A B

Now suppose $A :=$ event that $X=1$, $B :=$ event that $Y=1$, and $C :=$ event that $Z=1$.

These events satisfy the preconditions of the problem...

1. $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$ by independence of coin flips.
2. $\mathbb{P}(A \cap C) = \mathbb{P}(C \mid A) \cdot \mathbb{P}(A) = \mathbb{P}(C \mid \{(1,0),(1,1)\}) \cdot \mathbb{P}(A) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$
   $= \mathbb{P}(C) \cdot \mathbb{P}(A)$ ✓
3. $\mathbb{P}(B \cap C) = \mathbb{P}(C \mid B) \cdot \mathbb{P}(B)$ by symmetry of B & A and 2.

However, $\mathbb{P}(A \cap B \cap C) \neq \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C) = (1/2)^3 = 1/8$ since $A \cap B \cap C$ is impossible because $A, B \Rightarrow C^c$ and thus $\mathbb{P}(A \cap B \cap C) = 0$.

9) Let $X :=$ value of a dice roll. Since $X \sim \text{Uniform}(1,6)$ we have $\mu = 3.5$ and $\sigma^2 = \frac{6^2-1}{12} = 35/12$

↳ we know that we can approximate $\bar{X}_n = S_n/n \sim N(\mu, \sigma^2/n)$ so...

$$\mathbb{P}(\bar{X}_{40} > 5) \sim N(3.5, 7/96)$$

$$\Rightarrow \mathbb{P}\left(\frac{\bar{X}_{40} - 3.5}{\sqrt{7/96}} > \frac{1.5}{\sqrt{7/96}}\right) \sim N(0,1)$$

So... $\mathbb{P}\left(\frac{\bar{X}_{40} - 3.5}{\sqrt{7/96}} > 5.55\right) = 1 - \Phi(5.55) \approx 0$ Thus, $\boxed{\mathbb{P}(\bar{X}_{40} > 5) \approx 0}$

10) a) Let $Y = |X|$. Then, we have $Y = \begin{cases} X & \text{for } X \geq 0 \\ -X & \text{for } X < 0 \end{cases}$ and $Y \in [0, \infty)$, $X \in \mathbb{R}$

We want to find $F_Y(y) = \mathbb{P}(Y \leq y)$. Consider the support of Y...

↳ for any $y \geq 0$, $\mathbb{P}(Y \leq y) = \mathbb{P}(-y \leq X \leq y)$ since $Y \leq y \Rightarrow |X| \leq y \Rightarrow -y \leq X \leq y$

$= \mathbb{P}(X \leq y) - \mathbb{P}(X < -y)$ note: when splitting, need to handle subtracting out probability mass at $-y$ for general CDFs

$= \mathbb{P}(X \leq y) - \mathbb{P}(X < -y)$

$= F_X(y) - \lim_{t \to -y^-} F_X(t)$

OR... $\boxed{F_Y(y) = F_X(y) - F_X(-y) + \mathbb{P}(X = -y) \text{ on } y \in [0, \infty) \text{ (0 all else)}}$

b) Let $Z = X^+$. Then we have $Z = \begin{cases} 0 & \text{for } X < 0 \\ X & \text{for } X \geq 0 \end{cases}$ for $z \in [0, \infty)$, $x \in \mathbb{R}$

We want to find $\mathbb{P}(Z \leq z)$. Consider the support of Z...

↳ for $z > 0$ we have $\mathbb{P}(Z \leq z) = \mathbb{P}(X \leq z) = F_X(z)$

for $z = 0$ we have $\mathbb{P}(Z \leq z) = \mathbb{P}(X \leq 0) = F_X(0)$

So... $\boxed{F_Z = \begin{cases} F_X(z) & \text{for } z > 0 \\ F_X(0) & \text{for } z = 0 \\ 0 & \text{otherwise} \end{cases}}$

11) a) Note that for a given realization of $\{X_i\}_{i=1}^n$, we can order them such that $X_a \leq X_b \leq \dots \leq X$, where $a, b, \dots \in \{1, \dots, n\}$. Then, if $X_k \leq X$, it must be the case that all $X_i < X_k$ satisfy $X_i < X$.

↳ Thus, suppose Y is an RV such that $Y = X_k$ where k is selected uniformly on $\{1, 2, \dots, n\}$. We know that

$$\mathbb{P}(Y \leq x) = \mathbb{P}(X_k \leq x) = \frac{\#\text{ of } X_i \text{ below threshold } x}{\#\text{ of } X_i} \begin{matrix} \leftarrow \mathbb{1}. \\ \leftarrow n \end{matrix}$$

$$= \frac{\sum_{i=1}^{n} \mathbb{1}_{\{X_i \leq x\}}}{n} \quad \square.$$

b) $\lim_{n \to \infty} F_n(x) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{X_i \leq x\}} = \lim_{n \to \infty} \frac{\#\text{ of } X_i \leq x}{n}$. As we take $n \to \infty$, # of $X_i \leq x$ will tend to $n \cdot \mathbb{P}(X_i \leq x)$ by LLN, so $\lim_{n \to \infty} F_n(x) = \mathbb{P}(X_i \leq x) \ \forall i = F(x)$ for each fixed x.

Question 12

In this exercise, we will use simulations to empirically understand some findings.

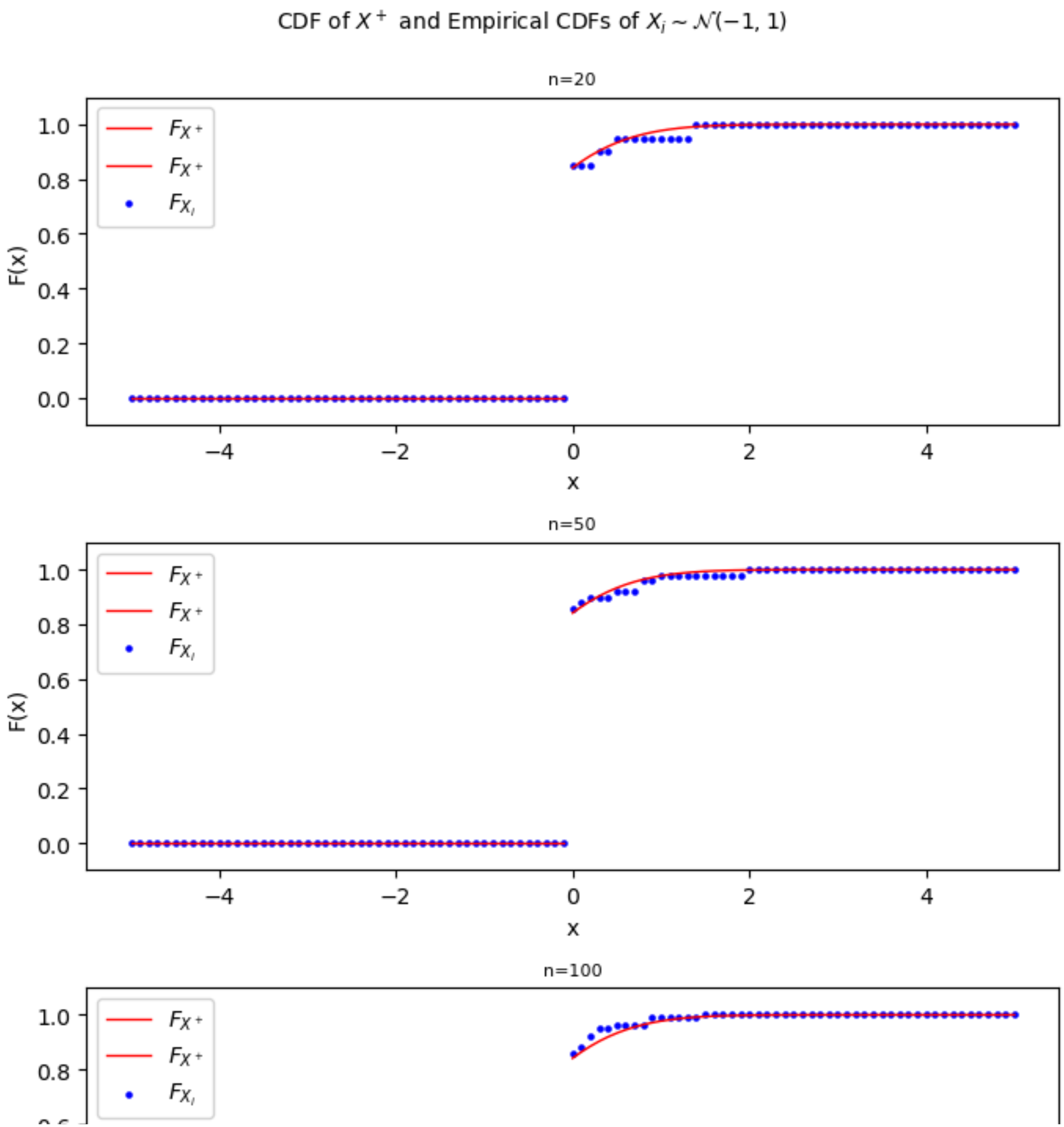
In [ ]:
```
# import statements
import numpy as np
import scipy as sp
import matplotlib.pyplot as plt
import math
```

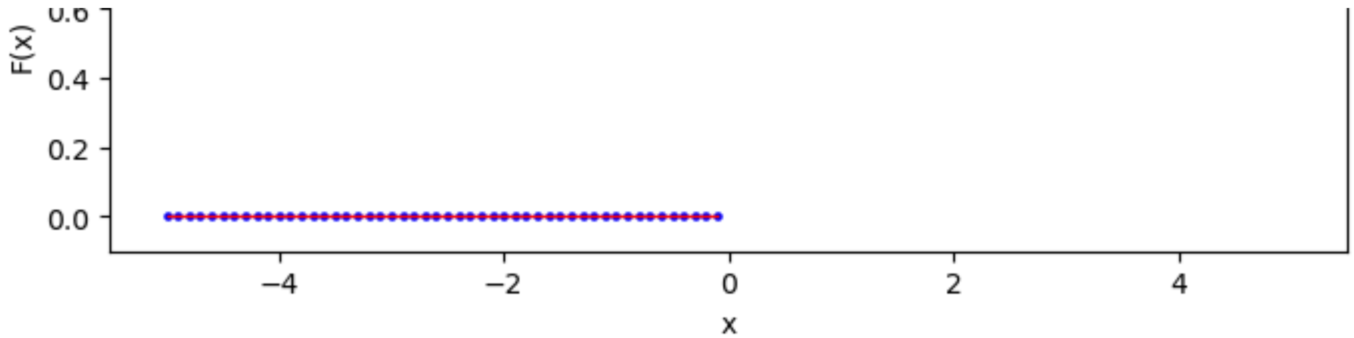
Part A

The goal here is to construct an approximation for the CDF of $X^+ = max\{0, X\}$ where $X \sim \mathcal{N}(-1, 1)$. We will simulate an n-sized sample $x_1, x_2, \ldots, x_n$, particularly with $n = 20$ from the aforementioned distribution.

In [155…
```
# plotting setup
fig, axes = plt.subplots(3, 1, figsize = (7,9))
fig.suptitle(r'CDF of $X^+$ and Empirical CDFs of $X_i \sim \mathcal{N}(-1,1)$', fontsize=10, y=0.99)

# create CDF approximation function of X+ = max{0,X}
def cdfApprox_xPlus(samples, x):
    return (1/(len(samples))) * sum([1 if max(s,0) <= x else 0 for s in samples])

# sample sizes to consider
n = [20, 50, 100]

# loop over different sample sizes
for i in range(3):
    # generate 20 samples from N(-1,1) using numpy random normal function
    samples = np.random.normal(-1,1,n[i])
    x = np.linspace(-5,5,101)

    # generate and plot CDF of X+ = max{0,X}
    pNeg = [0 for x in x if x < 0]
    pPos = [sp.stats.norm.cdf(0,-1,1)] + [sp.stats.norm.cdf(x,-1,1) for x in x if x > 0]
```

```python
    axes[i].plot([xi for xi in x if xi < 0],pNeg, color='red', label=r'$F_{X^+}$', linewidth=1)
    axes[i].plot([xi for xi in x if xi >= 0],pPos, color='red', label=r'$F_{X^+}$', linewidth=1)

    # plot CDF
    axes[i].scatter(x, [cdfApprox_xPlus(samples, xi) for xi in x], color='b', label=r'${F}_{X_i}$', s=5)

    # add legend and show plot
    axes[i].legend()
    axes[i].set_ylim(-0.1, 1.1)
    axes[i].set_xlabel('x')
    axes[i].set_ylabel('F(x)')
    axes[i].set_title(f'n={n[i]}', fontsize=8)

plt.tight_layout()
plt.show()
```


CDF of X^+ and Empirical CDFs of $X_i \sim \mathcal{N}(-1, 1)$
n=20
n=50
n=100
F_{X^+}
F_{X^+}
F_{X_i}
F(x)
x
1.0
0.8
0.6
0.4
0.2
0.0
−4
−2
0
2
4



Commentary

As expected, the derived CDF for X^+ is 0 for any $x < 0$, then jumps to ≈ 0.8 for $x = 0$ and proceeds as a continuous function.

We can see that the empirical CDFs (CDFs generated with n samples) are step functions, which makes sense since we are taking the average of n indicator variables given some threshold x. As x varies significantly, the average will likely change. However, there may be multiple x-values that have the same indicator average because the same subset of indicator variables are less than that set of x's. In other words, the number of indicator variables that satisfy the condition does not change until the condition is adequately raised or lowered, so the step function is an intuitive result.

Further, the empirical CDF approximation becomes increasingly smooth as n increases. It also converges to F pointwise as n increases, which illustrates the law of large numbers that we use to prove $F_n \to F$ as $n \to \infty$.

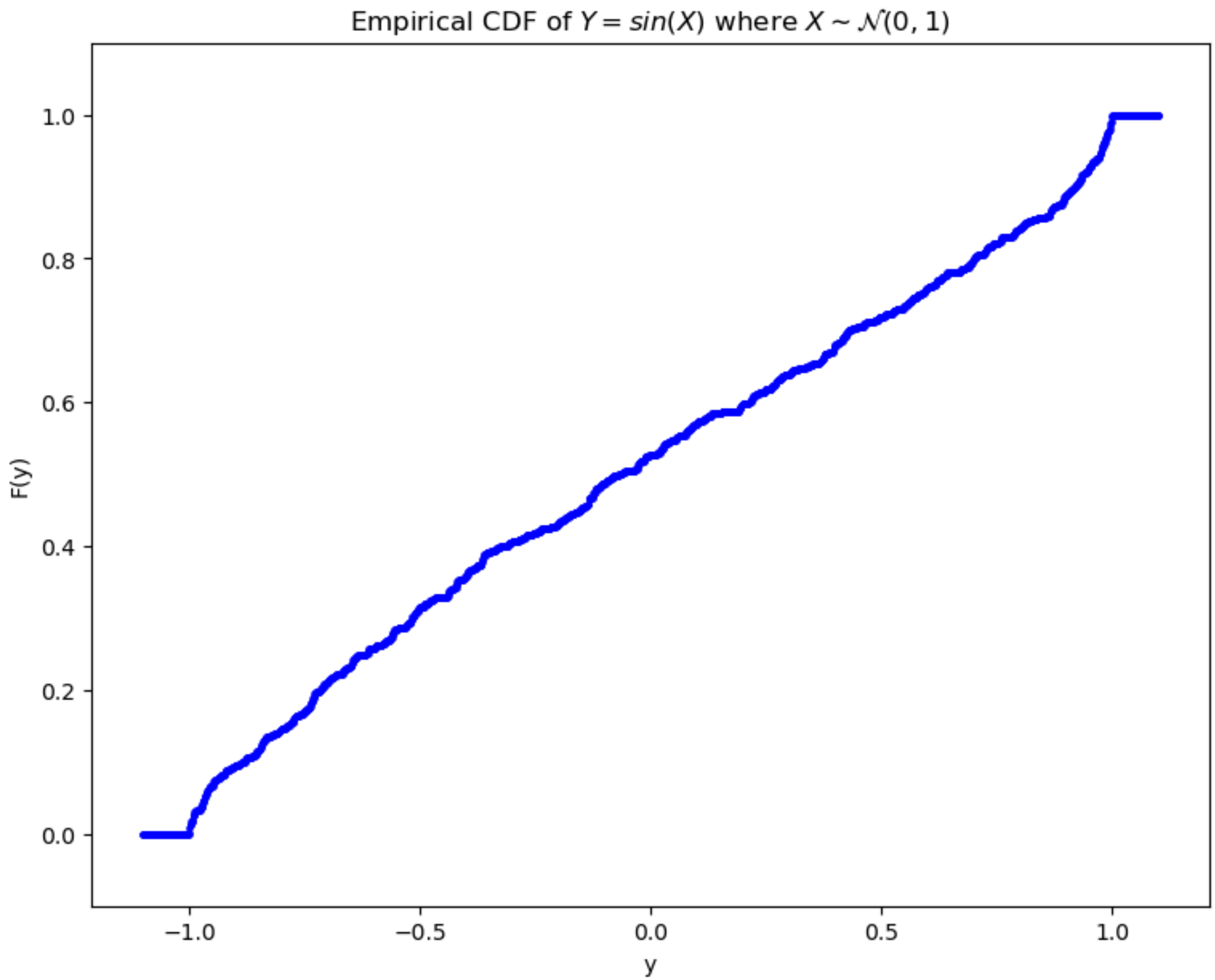
Part B

Let $X \sim \mathcal{N}(0,1)$ and $Y = sin(X)$. We will use simulations to draw an approximation of F_Y, the CDF of Y. Consider the mathematical derivation of this CDF.

In [157…

```
# plotting setup
plt.figure(figsize=(9,7))
plt.title(r'Empirical CDF of $Y = sin(X)$ where $X \sim \mathcal{N}(0,1)$')

# draw samples from N(0,1)
samples = np.random.normal(0,1,500)
```

```python
# create CDF approximation function of Y = sin(X)
def cdfApprox_sinY(samples, x):
    return (1/(len(samples))) * sum([1 if math.sin(s) <= x else 0 for s in samples])

# plot CDF
x = np.linspace(-1.1, 1.1, 2000)
plt.scatter(x, [cdfApprox_sinY(samples, xi) for xi in x], color='b', label=r'${F}_{Y}$', s=5)
plt.ylim(-0.1, 1.1)
plt.xlabel('y')
plt.ylabel('F(y)')
plt.show()
```


Empirical CDF of $Y = sin(X)$ where $X \sim \mathcal{N}(0, 1)$
1.0
0.8
0.6
0.4
0.2
0.0
F(y)
−1.0
−0.5
0.0
0.5
1.0
y