STA510
MANDATORY 1

Fall semester 2020

Student number 236227

Problem 1

Random variable X with pdf given by

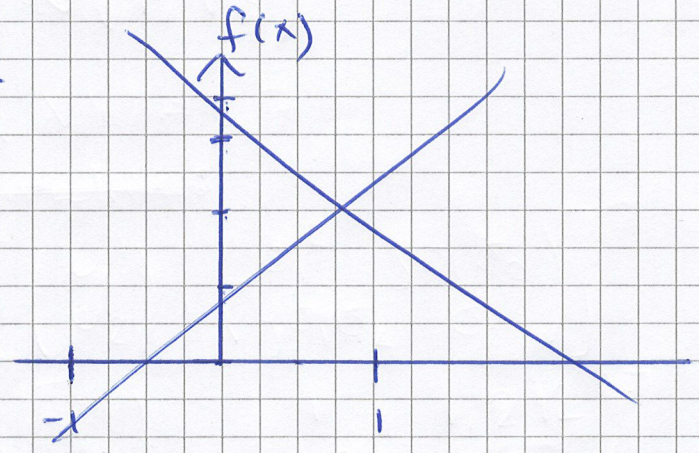
$$f(x) = \begin{cases} \frac{1}{2} - \frac{x}{4}, & -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

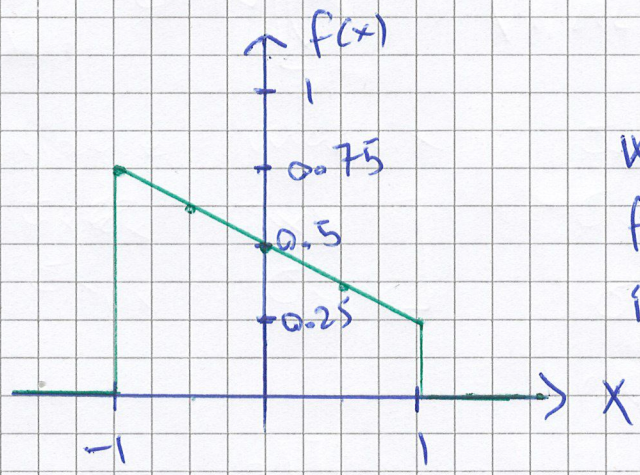
a) If this is a proper pdf, then

1. $f(x) \geq 0$ for all $x \in R$
2. $\int_{-\infty}^{\infty} f(x)dx = 1$
3. $P(a < X < b) = \int_a^b f(x)\,dx.$

Start by confirming number 1 with a plot:

x	$f(x) = \frac{1}{2} - \frac{x}{4}$
-1	0.75
-0.5	0.625
0	0.5
0.5	0.375
1	0.25





We see that $f(x) \geq 0$ for all $x \in R$, so point 1 is OK.

Now lets ~~ch~~ check point 2.

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-1}^{1} \frac{1}{2} - \frac{x}{4}\,dx$$

$$= \frac{1}{2}x - \frac{1}{8}x^2 \Big|_{x=-1}^{x=1} = \left(\frac{1}{2} - \frac{1}{8}\right) - \left(\frac{(-1)}{2} - \frac{(-1)^2}{8}\right) = 1$$

So we see that

$\int_{-\infty}^{\infty} f(x)\,dx = 1$, which confirmes point 2.

Point 3, we can calculate

$P(a < X < b)$ by $\int_a^b f(x)\,dx$

This was proved in point 1 by $\int_{-1}^{1} f(x)\,dx$, which is the probability $P(-1 < x < 1)$, but we may as well prove it for another case, such as

$$P(0 < x < 1) = \int_0^1 f(x)\,dx = \frac{1}{2}x - \frac{x^2}{8}\Big|_0^1 = \frac{3}{8}.$$

This corresponds to the area under the curve between 0 and 1.

<u>Therfore, f(x) is a proper probability density function.</u>

By definition, $E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$

$$\underline{\underline{E(X)}} = \int_{-1}^{1} x\left(\frac{1}{2} - \frac{x}{4}\right) dx = \int_{-1}^{1} \frac{x}{2} - \frac{x^2}{4}\, dx$$

$$= \frac{x^2}{4} - \frac{x^3}{12}\Big|_{-1}^{1} = \left(\frac{1^2}{4} - \frac{1^3}{12}\right) - \left(\frac{(-1)^2}{4} - \frac{(-1)^3}{12}\right)$$

$$= \underline{\underline{\frac{-1}{6}}}$$

By definition, $Var(X) = \int_{-\infty}^{\infty} (x-\mu) f(x)\, dx$, where $\mu = E(X)$. This is equivalent with:

$Var(X) = E(X^2) - E(X)^2$.

$$E(X^2) = \int_{-1}^{1} x^2\left(\frac{1}{2} - \frac{x}{4}\right) dx = \int_{-1}^{1} \frac{x^2}{2} - \frac{x^3}{4}\, dx$$

$$= \frac{x^3}{6} - \frac{x^4}{16}\Big|_{-1}^{1} = \left(\frac{1^3}{6} - \frac{1^4}{16}\right) - \left(\frac{(-1)^3}{6} - \frac{(-1)^4}{16}\right)$$

$$= \frac{1}{3}$$

$$\underline{\underline{Var(X)}} = E(X^2) - E(X)^2 = \frac{1}{3} - \left(-\frac{1}{6}\right)^2 = \underline{\underline{\frac{11}{36}}} \quad (\approx 0.3056)$$

$$E\left(\frac{1}{2}(X+1)\right) = E\left(\frac{1}{2}X + \frac{1}{2}\right) = \frac{1}{2}E(X) + \frac{1}{2}$$

$$= \frac{1}{2}\cdot\left(\frac{-1}{6}\right) + \frac{1}{2}$$

$$\underline{\underline{E\left(\frac{1}{2}(X+1)\right) = \frac{5}{12}}}$$

$$E(X^3) = \int_{-1}^{1} x^3 f(x)\,dx = \int_{-1}^{1} x^3\left(\frac{1}{2} - \frac{x}{4}\right)dx$$

$$= \int_{-1}^{1} \frac{x^3}{2} - \frac{x^4}{4}\,dx = \frac{x^4}{8} - \frac{x^5}{20}\Big|_{-1}^{1}$$

$$= \left(\frac{1^4}{8} - \frac{1^5}{20}\right) - \left(\frac{(-1)^4}{8} - \frac{(-1)^5}{20}\right)$$

$$\underline{\underline{E(X^3) = -\frac{1}{10}}}$$

b) By definition, cdf of a continous random variable X is

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

$$F(x) = P(X \leq x) = \int_{-1}^{x} f(t)dt = \int_{-1}^{x} \frac{1}{2} - \frac{t}{4} dt$$

$$= \frac{1}{2}t - \frac{1}{8}t^2 \Big|_{-1}^{x} = \left(\frac{1}{2}x - \frac{1}{8}x^2\right) - \left(\frac{(-1)}{2} - \frac{(-1)^2}{8}\right)$$

$$\underline{\underline{F(x)}} = \frac{x}{2} - \frac{x^2}{8} + \frac{5}{8} = \underline{\underline{\frac{5}{8} + \frac{x}{2}\left(1 - \frac{x}{4}\right)}}$$

$$\underline{\underline{P(X < 0)}} = \frac{5}{8} + \frac{0}{2}\left(1 - \frac{0}{4}\right) = \underline{\underline{\frac{5}{8}}}$$

$$\underline{\underline{P(-0.5 < X < 0.5)}} = P(X < 0.5) - (1 - P(X < -0.5))$$

$$= \frac{5}{8} + \frac{0.5}{2}\left(1 - \frac{0.5}{4}\right) - \left(1 - \left(\frac{5}{8} + \frac{(-0.5)}{2}\left(1 - \frac{(-0.5)}{4}\right)\right)\right)$$

$$= \underline{\underline{\frac{3}{16}}}$$

To find the inverse cumulative distribution function, I set $u = F(x)$ and solve for x:

$$u = F(x) = \frac{5}{8} + \frac{x}{2}\left(1 - \frac{x}{4}\right) = \frac{5}{8} + \frac{x}{2} - \frac{x^2}{8}$$

$$8u = 5 + 4x - x^2$$

$$0 = -x^2 + 4x + 5 - 8u$$

This is quadratic equation with $a = -1$, $b = 4$, $c = 5 - 8u$ which yields $x_1 = 2 + \sqrt{-8u+9}$ and $x_2 = 2 - \sqrt{-8u+9}$.

Only $x_2 = 2 - \sqrt{-8u+9}$ is a sensical answer, since x_1 is out of range (undefined)

So we get

$$0 = x - 2 + \sqrt{-8u+9}$$

$$x = \underline{\underline{2 - \sqrt{-8u+9} = F^{-1}(u)}}$$

d) Acceptance-rejection method.

We have $f(x) = \frac{1}{2} - \frac{x}{4}$ for $-1 \le x \le 1$.

Let $g(x)$ be uniform$(-1,1)$.

We know that $f(x)$ is at a maximum value when $x = -1$

$$f(-1) = \frac{1}{2} - \frac{(-1)}{4} = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$$

Then $\dfrac{f(x)}{g(x)} = \dfrac{\frac{3}{4}}{\frac{1}{2}} \le \frac{3}{2}$ for all $-1 \le x \le 1$,

so $c = 3/2$

A random x from $g(x)$ is accepted if

$$\cancel{\frac{f(x)}{c\,g(x)} = \frac{\frac{1}{2} - \frac{x}{4}}{\frac{3}{2} \cdot \frac{1}{2}} = \frac{\frac{1}{2} - \frac{x}{4}}{\frac{3}{4}} = \frac{2-x}{4} \cdot \frac{4}{3}}$$

$$\cancel{= \frac{4 - 2x}{3}}$$

$$\frac{f(x)}{c \cdot g(x)} = \frac{\frac{1}{2} - \frac{x}{4}}{\frac{3}{2} \cdot \frac{1}{2}} = \frac{2-x}{\cancel{4}} \cdot \frac{\cancel{4}}{3} = \underline{\underline{\frac{2-x}{3} > u}}$$

Problem 2

$X_1, X_2, \ldots, X_n$ are i.i.d $N(\mu, \sigma^2)$

a)

$$A = \sum_{i=1}^{n} X_i$$

$$E(A) = E\left(\sum_{i=1}^{n} X_i\right) = E(X_1) + E(X_2) + \ldots + E(X_n) = n\mu$$

$$Var(A) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) = n\sigma^2$$

$$\underline{\underline{A \sim N(n\mu, n\sigma^2)}}$$

$$B = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$E(B) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n} \cdot n\mu = \mu$$

$$Var(B) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \left(\frac{1}{n}\right)^2 Var(X_1) + \ldots + \left(\frac{1}{n}\right)^2 Var(X_n)$$

$$= \left(\frac{1}{n}\right)^2 \sigma^2 + \ldots + \left(\frac{1}{n}\right)^2 \sigma^2$$

$$= \frac{\sigma^2}{n}$$

$$\underline{\underline{B \sim N\left(\mu, \frac{\sigma^2}{n}\right)}}$$

$$C = \frac{\sqrt{n}(B-\mu)}{\sigma}, \quad B = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$C = \frac{\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right)}{\sigma} = \frac{\frac{\sqrt{n}}{n}\sum_{i=1}^{n} X_i - \sqrt{n}\,\mu}{\sigma}$$

$$C = \frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i}{\sigma} - \frac{\sqrt{n}\,\mu}{\sigma}$$

$$\frac{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i}{\sigma} = \frac{\frac{\sum_{i=1}^{n} X_i}{\sqrt{n}}}{\frac{\sigma}{1}} = \frac{\sum_{i=1}^{n} X_i}{\sqrt{n}} \cdot \frac{1}{\sigma} = \frac{\sum_{i=1}^{n} X_i}{\sqrt{n}\sigma} = \frac{1}{\sqrt{n}\sigma} \cdot \sum_{i=1}^{n} X_i$$

$$C = \frac{1}{\sqrt{n}\sigma} \cdot \sum_{i=1}^{n} X_i - \frac{\sqrt{n}\,\mu}{\sigma}$$

$$= \sum_{i=1}^{n} a_i X_i + b, \text{ where}$$

$$a_i = \frac{1}{\sqrt{n}\sigma} \text{ and } b = -\frac{\sqrt{n}\,\mu}{\sigma}$$

$$E(C) = \underbrace{\frac{1}{\sqrt{n}\sigma} \cdot \mu + \dots + \frac{1}{\sqrt{n}\sigma} \cdot \mu}_{n \text{ times}} - \frac{\sqrt{n}\,\mu}{\sigma}$$

$$= \frac{n\mu}{\sqrt{n}\sigma} - \frac{\sqrt{n}\,\mu}{\sigma} = \frac{\sqrt{n}\,\mu}{\sigma} - \frac{\sqrt{n}\,\mu}{\sigma} = 0$$

$$Var(C) = \sum_{i=1}^{n} \left(\frac{1}{\sqrt{n}\sigma}\right)^2 \cdot Var(X) = \left(\frac{1}{n\sigma^2}\right)\sigma^2 + \dots + \left(\frac{1}{n\sigma^2}\right)\sigma^2 = 1$$

$$\underline{\underline{C \sim N(0,1)}}$$

c) 95% Confinence interval is no longer than 0.1

$$n > \frac{4\sigma^2}{e^2}$$ (from lecture notes "chap2_part2")

Here, $\sigma^2 = 2^2$ and $e^2 = (1-0.95)^2 = (0.05)^2$

$$n > \frac{4 \cdot 2^2}{0.05^2} = 6400$$

<u>So we need at least 6400 simulations</u>

Sources as used in the .R file

1. *"quantiles"*, rdocumentation.org. Accessed 16.09.20 [online]. Available at: https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/quantile
2. *"cor"*, rdocumentation.org. Accessed 16.09.20 [online]. Available at: https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/cor
3. *"Pearson correlation - An overview"*, Science direct topics. Accessed 16.09.20 [online]. Available at: https://www.sciencedirect.com/topics/computer-science/pearson-correlation