Question 1 (suggested for peer/personal tutorial)

Recall from Section 8.3.8 in Prof. Veraart's notes that the p.d.f. of the standard Cauchy distribution is

$$f_X(x) = \frac{1}{\pi\left(1+x^2\right)}, \qquad \text{with support } x \in \mathbb{R}.$$

(a) Show that f_X is a probability density function (p.d.f.) and plot $y = f_X(x)$.

(b) Compute the first (raw) moment of X, $\mu = \mu_1' = \mathrm{E}(X)$.

(c) Compute the kth central moment of X, $\mu_k = \mathrm{E}\left((X-\mu)^k\right)$ for $k \in \{2, 3, \ldots\}$.

(d) Compute the second raw moment of X, $\mu_2' = \mathrm{E}\left(X^2\right)$.

Solution to Question 1 (suggested for peer/personal tutorial)

Part (a): For f_X to be a p.d.f., it needs to (i) be nonnegative on $\mathbb{R}$ and (ii) integrate to 1. Part (i) is true since $x^2 \geq 0$ for all real x, and therefore $f_X(x) \geq 0$ for all real x. For (ii), recall that

$$\frac{\mathrm{d}}{\mathrm{d}x} \arctan x = \frac{1}{1+x^2}.$$

and

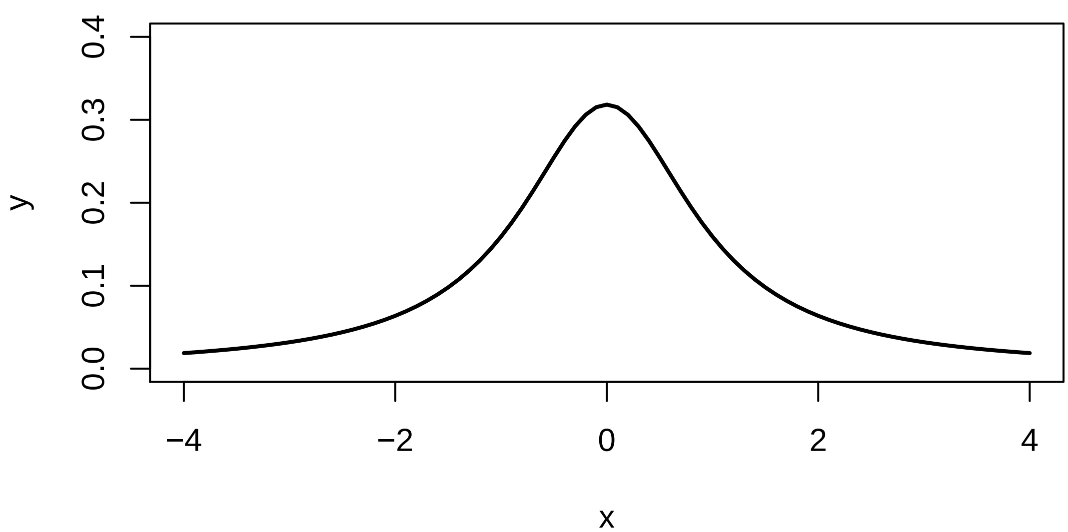
$$\lim_{\theta \to \frac{\pi}{2}} \tan\theta = \infty, \qquad \lim_{\theta \to -\frac{\pi}{2}} \tan\theta = -\infty.$$

Then,

$$\int_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1}{(1+x^2)}\,\mathrm{d}x = \frac{1}{\pi}\left[\arctan x\right]_{-\infty}^{\infty} = \frac{1}{\pi}\left[\tfrac{\pi}{2} - \left(-\tfrac{\pi}{2}\right)\right] = 1.$$

One can take the first derivative to show that show that f_X has a local maximum at $(0, \frac{1}{\pi}) = (0, 0.318)$, and notice that $\lim_{x\to\infty} f_X(x) = \lim_{x\to-\infty} f_X(x) = 0$ to plot:

```r
x <- seq(from=-4, to=4, by=0.1)
y <- dcauchy(x)
plot(x, y, type='l', lwd=2, ylim=c(0, 0.4))
```



One could compute the second derivative to obtain the two inflection points $\left(\pm\frac{1}{\sqrt{3}}, \frac{3}{4\pi}\right) = (\pm 0.577, 0.239)$, but it is not essential to this exercise.

In order to see $\frac{\mathrm{d}}{\mathrm{d}x}\arctan x = \frac{1}{1+x^2}$, first recall the trigonemtric identity:

$$\sin^2\theta + \cos^2\theta = 1 \qquad \Rightarrow \frac{\sin^2\theta}{\cos^2\theta} + \frac{\cos^2\theta}{\cos^2\theta} = \frac{1}{\cos^2\theta} \qquad \Rightarrow \tan^2\theta + 1 = \sec^2\theta$$

And using implicit differentiation,

$$y = \arctan x \qquad \Rightarrow \tan y = x \qquad \Rightarrow \frac{\mathrm{d}}{\mathrm{d}x}\tan y = \frac{\mathrm{d}}{\mathrm{d}x}x \qquad \Rightarrow \sec^2 y\frac{\mathrm{d}y}{\mathrm{d}x} = 1 \qquad \Rightarrow \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{1+x^2}.$$

To find the derivative of $\tan\theta$, write $\tan\theta = \frac{\sin\theta}{\cos\theta}$ and use the quotient rule.

Part (b): Looking at the plot of the p.d.f., one would guess that $\mathrm{E}(X) = 0$. However, when one does the computation:

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x}{1+x^2}\,\mathrm{d}x = \frac{1}{\pi}\cdot\frac{1}{2}\left[\log\left(1+x^2\right)\right]_{-\infty}^{\infty}.$$

At this point it is tempting to say that since $\lim_{x\to\infty}\log(1+x^2) = \infty$ and $\lim_{x\to-\infty}\log(1+x^2) = \infty$ and so the two terms "cancel out" and the answer is 0, **but this is incorrect**.

Remark 1.1.18 in the lecture notes recalled that when computing moments and dealing with improper integrals, one must consider the two partial integrals in the right hand side of

$$\int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x = \int_{-\infty}^{0} x f_X(x)\,\mathrm{d}x + \int_{0}^{\infty} x f_X(x)\,\mathrm{d}x,$$

(where for this example we split the integral at $a = 0$). Then in order for the integral on the left-hand side to exist, either:

- at least one of the integrals on the right-hand converges, or
- both integrals are ∞, or both are $-\infty$.

The reason for the second point is that we cannot have the situation where one term is ∞ and the other is $-\infty$; in that case the integral on the left-hand side ("$\infty - \infty$") is undefined.

However, this is precisely the situation we now have:

$$\int_{-\infty}^{0} x f_X(x)\,\mathrm{d}x = \frac{1}{\pi}\cdot\frac{1}{2}\left[\log\left(1+x^2\right)\right]_{-\infty}^{0} = -\infty,$$

$$\int_{0}^{\infty} x f_X(x)\,\mathrm{d}x = \frac{1}{\pi}\cdot\frac{1}{2}\left[\log\left(1+x^2\right)\right]_{0}^{\infty} = \infty.$$

Therefore, **for the Cauchy distribution, $\mu = \mathrm{E}(X)$ is undefined**.

Part (c): For $k \in \{2, 3, \ldots\}$, the computation for the kth **central** moment involves adding/subtracting (a power of) the first moment μ, which is shown in Part (b) to be undefined. Therefore, the central moments are all undefined (for $k \geq 2$).

Part (d): Although the first moment is undefined, the second central moment is defined, but infinite:

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x^2}{1+x^2}\,\mathrm{d}x = \frac{1}{\pi}\int_{-\infty}^{\infty}\left[1 - \frac{1}{1+x^2}\right]\mathrm{d}x = \frac{1}{\pi}\left[x - \arctan x\right]_{-\infty}^{\infty} = \frac{1}{\pi}\left[x\right]_{-\infty}^{\infty} - \frac{1}{\pi}\left[\arctan x\right]_{-\infty}^{\infty} = \infty - 1 = \infty.$$

Notice the difference between a moment being undefined and a moment being infinite.

Question 2

Suppose that the random variable X is known to only take non-zero values in the bounded range $[a, b]$, i.e. the support of X is $[a, b]$.

(a) Prove that $\mathrm{Var}\,(X) \leq \dfrac{(b-a)^2}{4}$.

(b) Conclude that if $X \sim \text{Bernoulli}(p)$, for some $p \in [0, 1]$, then $\mathrm{Var}\,(X) \leq \frac{1}{4}$.

(c) Is the bound $\mathrm{Var}\,(X) \leq \dfrac{(b-a)^2}{4}$ tight? In other words, is there a distribution F_X with support $[a, b]$ for which $\mathrm{Var}\,(X) = \dfrac{(b-a)^2}{4}$?

Solution to Question 2

Part (a): We first derive the expression $(X-a)(X-b) = \left(X - \frac{a+b}{2}\right)^2 - \frac{(b-a)^2}{4}$ by completing the square:

$$\begin{aligned}
(X-a)(X-b) &= X^2 - (a+b)X + ab \\
&= \left(X^2 - (a+b)X + \left(\tfrac{a+b}{2}\right)^2\right) + ab - \tfrac{(a+b)^2}{4} \\
&= \left(X - \tfrac{a+b}{2}\right)^2 + \tfrac{1}{4}\left[4ab - (a+b)^2\right] \\
&= \left(X - \tfrac{a+b}{2}\right)^2 + \tfrac{1}{4}\left[4ab - (a^2 + 2ab + b^2)\right] \\
&= \left(X - \tfrac{a+b}{2}\right)^2 + \tfrac{1}{4}\left[-(a^2 - 2ab + b^2)\right] \\
&= \left(X - \tfrac{a+b}{2}\right)^2 - \tfrac{1}{4}\left[(a-b)^2\right] \\
&= \left(X - \tfrac{a+b}{2}\right)^2 - \tfrac{(b-a)^2}{4}.
\end{aligned}$$

Note this implies that

$$\left(X - \tfrac{a+b}{2}\right)^2 = (X-a)(X-b) + \tfrac{(b-a)^2}{4}. \tag{1}$$

Recall Theorem 1.1.1: Given a random variable X, then over all values $c \in \mathbb{R}$,

$$\min_c \mathrm{E}[(X-c)^2] = \mathrm{E}[(X - \mathrm{E}[X])^2].$$

This implies that for any $c \in \mathbb{R}$,

$$\mathrm{Var}\,(X) = \mathrm{E}[(X - \mathrm{E}[X])^2] \leq \mathrm{E}[(X-c)^2].$$

In particular, using $c = \frac{a+b}{2}$, and Equation (1),

$$\mathrm{Var}\,(X) \leq \mathrm{E}[\left(X - \tfrac{a+b}{2}\right)^2] = \mathrm{E}[(X-a)(X-b) + \tfrac{(b-a)^2}{4}] = \mathrm{E}[(X-a)(X-b)] + \tfrac{(b-a)^2}{4}$$

since $\frac{(b-a)^2}{4}$ is a constant, and using the linearity of the expectation. Now, since $X \in [a, b]$, this implies

$$\begin{aligned}
&(X-a)(X-b) \leq 0, \\
\Rightarrow &\mathrm{E}[(X-a)(X-b)] \leq 0,
\end{aligned}$$

and therefore

$$\mathrm{Var}\,(X) \leq \mathrm{E}[(X-a)(X-b)] + \tfrac{(b-a)^2}{4} \leq \tfrac{(b-a)^2}{4}.$$

This proves Part (a). For **Part (b)**, since any Bernoulli random variable X only takes values of either 0 or 1, it itself is bounded in the range $[0, 1]$, and the result follows. (Note: it has nothing to do with the value of p.)

For **Part (c)**, consider the discrete distribution F_X for the random variable X defined on $[a, b]$ by

$$X = \begin{cases} a, & \text{with probability } \frac{1}{2} \\ b, & \text{with probability } \frac{1}{2} \end{cases}.$$

Then

$$\mathrm{E}\left(X\right) = \sum_x x\mathrm{P}\left(X = x\right) = a \cdot \mathrm{P}\left(X = a\right) + b \cdot \mathrm{P}\left(X = b\right) = \frac{a+b}{2}$$
$$\mathrm{E}\left(X^2\right) = \sum_x x^2\mathrm{P}\left(X = x\right) = a^2 \cdot \mathrm{P}\left(X = a\right) + b^2 \cdot \mathrm{P}\left(X = b\right) = \frac{a^2+b^2}{2}$$
$$\mathrm{Var}\left(X^2\right) = \mathrm{E}\left(X^2\right) - \left(\mathrm{E}\left(X\right)\right)^2 = \frac{a^2+b^2}{2} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{4}$$

Therefore, the bound on the variance is tight, since this probability distribution has exactly this variance.

One might wonder what would inspire one to think of this example. First, one has to remember to work within the constraint that the support is $[a, b]$. Then, in order to maximise the variance, the idea (might) be to separate the probability mass function to place half of the mass at one extreme and the other half of the mass at the other extreme. The resulting thinking would lead one to the discrete distribution defined above. Of course, this is only an outline of one possible train of thought that would lead one to this solution.

Question 3

Suppose that $X_1, X_2, \ldots, X_n$ are independent random variables all following a $\mathrm{N}(0, 1)$ distribution.

(a) What is the distribution of the random variable $Z = X_1^2$?

(b) Find the distribution of $Y = X_1^2 + X_2^2 + \cdots + X_n^2$.

Solution to Question 3

Part (a): If $X_1 \sim \mathrm{N}(0, 1)$, and $Z = X_1^2$, then $Z \sim \chi_1^2$, a chi-squared random variable with one degree of freedom, which has p.d.f. and support

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} z^{-1/2} e^{-z/2}, \qquad z > 0.$$

To see the derivation of this result, please see the solution to Problem Sheet 1, Exercise 8 (Term 1). It is worth familiarising yourself with this derivation if you do not remember it.

Also it is worth recalling from Section 8.3.4 of Prof. Veraart's lecture notes that the p.d.f. of the χ_n^2 for $n = 2, 3, \ldots$ is

$$f_W(w) = \frac{1}{2^{n/2}\Gamma(n/2)} w^{n/2-1} e^{-w/2}, \qquad z \geq 0.$$

Part (b): The easiest way to do this part is to use moment generating functions (m.g.f.'s). The m.g.f. of a χ_1^2 random variable is

$$M_Z(t) = \mathrm{E}_Z\left(e^{tz}\right) = \int_0^\infty e^{tz} \frac{1}{\sqrt{2\pi}} z^{-1/2} e^{-z/2}\, \mathrm{d}z = \int_0^\infty \frac{1}{\sqrt{2\pi}} z^{-1/2} e^{-z(1-2t)/2}\, \mathrm{d}z$$

If we use the substitution $u = z(1-2t)$, and constrain $t < \frac{1}{2}$,

$$\begin{aligned} M_Z(t) &= \int_0^\infty \frac{1}{\sqrt{2\pi}} \left(\frac{u}{1-2t}\right)^{-1/2} e^{-u/2} \left(\frac{1}{1-2t}\right) \mathrm{d}u \\ &= \left(\frac{1}{1-2t}\right)^{-1/2} \left(\frac{1}{1-2t}\right) \int_0^\infty \frac{1}{\sqrt{2\pi}} u^{-1/2} e^{-u/2} \, \mathrm{d}u \\ &= (1-2t)^{-1/2} \end{aligned}$$

since the integral on the right is 1, because the integrand is the p.d.f. is of a χ^2_1 random variable (see above). Similarly, using the p.d.f. above one can show that the m.g.f. of a χ^2_n random variable (for $n \in \{1, 2, \ldots\}$) is

$$M_W(t) = (1-2t)^{-n/2}, \qquad t < \frac{1}{2}.$$

Now, one can use the Theorem 12.2.7 in Prof. Veraart's notes. Since the random variables $X_1, X_2, \ldots, X_n$ are independent, their squares $X_1^2, X_2^2, \ldots, X_n^2$ are also independent, and so defining $Y = X_1^2 + X_2^2 + \cdots + X_n^2$,

$$M_Y(t) = M_{\sum_{i=1}^n X_i^2}(t) = \prod_{i=1}^n M_{X_i^2}(t) = \prod_{i=1}^n (1-2t)^{-1/2} = (1-2t)^{-n/2} = M_W(t),$$

which shows that Y has a χ^2_n distribution.

Question 4

Complete Exercise 1.2.5 in the lecture notes: Given a sample of observations $x_1, x_2, \ldots, x_n$, with the sample mean $\overline{x}$ and sample variance s^2 defined in Definition 1.2.9 of the lecture notes, prove that

$$\min_a \left[\sum_{i=1}^n (x_i - a)^2\right] = \sum_{i=1}^n (x_i - \overline{x})^2 = (n-1)\, s^2.$$

Solution to Question 4

We use a similar trick to that in the proof of Theorem 1.1.1. For any given a,

$$\begin{aligned} \sum_{i=1}^n (x_i - a)^2 = \sum_{i=1}^n \left[(x_i - \overline{x}) + (\overline{x} - a)\right]^2 &= \sum_{i=1}^n \left[(x_i - \overline{x})^2 + 2\,(x_i - \overline{x})\,(\overline{x} - a) + (\overline{x} - a)^2\right] \\ &= \sum_{i=1}^n (x_i - \overline{x})^2 + 2\,(\overline{x} - a) \sum_{i=1}^n (x_i - \overline{x}) + \sum_{i=1}^n (\overline{x} - a)^2 \\ &= \sum_{i=1}^n (x_i - \overline{x})^2 + 2\,(\overline{x} - a) \cdot 0 + n\,(\overline{x} - a)^2 \\ &= \sum_{i=1}^n (x_i - \overline{x})^2 + n\,(\overline{x} - a)^2, \end{aligned}$$

where in the third line we used

$$\sum_{i=1}^n (x_i - \overline{x}) = \sum_{i=1}^n x_i - \sum_{i=1}^n \overline{x} = n\overline{x} - n\overline{x} = 0.$$

Since $n\,(\overline{x} - a)^2 \geq 0$,

$$\sum_{i=1}^n (x_i - a)^2 \leq \sum_{i=1}^n (x_i - \overline{x})^2,$$

with equality only when $\overline{x} = a$, which proves the result.

Question 5 (using R)

(a) Use R to generate 10 observations from a normal distribution with mean 3 and variance 2. Save the values in a vector `x`.

(b) Use the built-in R commands to compute the sample mean, variance and standard deviation of `x`.

(c) Write your own R functions to compute the sample mean and sample variance of `x`.

Solution to Question 5

Part (a):

```
set.seed(1)
x <- rnorm(n=10, mean=3, sd = sqrt(2))
print(x)
#>  [1] 2.114 3.260 1.818 5.256 3.466 1.840 3.689 4.044 3.814
#> [10] 2.568
```

Part (b):

```
print(mean(x))
#> [1] 3.187

print(var(x))
#> [1] 1.219

print(sd(x))
#> [1] 1.104
```

Part (c): It is also possible to use for-loops, but the built-in function `sum` is quite useful:

```
my_mean <- function(x){
    return(  sum(x) / length(x)  )
}

my_var <- function(x){
    return(   sum(  ( x - mean(x) )^2  )  /  (length(x) - 1)   )
}

print(my_mean(x))
#> [1] 3.187

print(my_var(x))
#> [1] 1.219
```