Exercise 1.  
a) i) Remembering that $e^\lambda = \sum_{k \geq 0} \frac{\lambda^k}{k!}$, we obtain

$$\sum_{k \geq 0} \frac{\lambda^k e^{-\lambda}}{k!} = e^\lambda e^{-\lambda} = 1$$

ii) Similarly, we obtain:

$$E(X) = \sum_{k \geq 0} k \frac{\lambda^k e^{-\lambda}}{k!} = \lambda \sum_{k \geq 1} \frac{\lambda^{k-1} e^{-\lambda}}{(k-1)!} = \lambda$$

iii) as well as:

$$E(X(X-1)) = \sum_{k \geq 0} k(k-1) \frac{\lambda^k e^{-\lambda}}{k!} = \lambda^2 \sum_{k \geq 2} \frac{\lambda^{k-2} e^{-\lambda}}{(k-2)!} = \lambda^2$$

iv)

$$\text{Var}(X) = E(X^2) - E(X)^2 = E(X(X-1)) + E(X) - E(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

b) i) Following the hint, we obtain

$$\int \int_{\mathbb{R}^2} dx dy p_X(x) p_X(y) = \frac{1}{2\pi \sigma^2} \int_0^\infty dr \int_0^{2\pi} d\theta \exp \left( -\frac{r^2}{2\sigma^2} \right) = \frac{1}{\sigma^2} \left( -\sigma^2 \exp \left( -\frac{r^2}{2\sigma^2} \right) \right) \bigg|_0^{+\infty} = 1$$

which proves the claim.

ii) First observe that because the mapping $x \mapsto x p_X(x)$ is odd, $E(X) = 0$. By integration by parts (noting that the boundary values at $\pm \infty$ vanish, because the exponential decay wins over any polynomial function), we obtain:

$$\text{Var}(X) = E(X^2) = \frac{1}{\sqrt{2\pi \sigma^2}} \int_\mathbb{R} x \cdot x \exp \left( -\frac{x^2}{2\sigma^2} \right) dx = \frac{1}{\sqrt{2\pi \sigma^2}} \int_\mathbb{R} 1 \cdot \sigma^2 \exp \left( -\frac{x^2}{2\sigma^2} \right) dx = \sigma^2$$

iii) Using again integration by parts, we obtain

$$E(X^4) = \frac{1}{\sqrt{2\pi \sigma^2}} \int_\mathbb{R} x^3 \cdot x \exp \left( -\frac{x^2}{2\sigma^2} \right) dx = \frac{1}{\sqrt{2\pi \sigma^2}} \int_\mathbb{R} 3x^2 \cdot \sigma^2 \exp \left( -\frac{x^2}{2\sigma^2} \right) dx = 3\sigma^4$$

iv)

$$E(\exp(X)) = \frac{1}{\sqrt{2\pi \sigma^2}} \int_\mathbb{R} \exp \left( x - \frac{x^2}{2\sigma^2} \right) dx = \frac{1}{\sqrt{2\pi \sigma^2}} \int_\mathbb{R} \exp \left( -\frac{(x-\sigma)^2}{2\sigma^2} + \frac{\sigma^2}{2} \right) dx = \exp \left( \frac{\sigma^2}{2} \right)$$

v) Note here that the expectation is always well defined (as $\exp(X^2)$ is a non-negative random variable), but can take the value $+\infty$, depending on the value of $\sigma$:

$$E(\exp(X^2)) = \frac{1}{\sqrt{2\pi \sigma^2}} \int_\mathbb{R} \exp \left( -x^2 \left( \frac{1}{2\sigma^2} - 1 \right) \right) dx = \begin{cases} +\infty & \text{if } \sigma \geq 1/\sqrt{2} \\ \frac{1}{\sqrt{2\pi \sigma^2}} \sqrt{\frac{\pi}{2\sigma^2} - 1} = \frac{1}{\sqrt{1-2\sigma^2}} & \text{if } \sigma < 1/\sqrt{2} \end{cases}$$
**Exercise 2.** a) Here is a generic procedure to generate a random variable $X$ with (strictly) increasing cdf $F_X$ from a uniform random variable $U \sim U([0,1])$. Observe first that the random variable $F_X^{-1}(X)$ has cdf

$$P(\{F_X(X) \leq t\}) = P(\{X \leq F_X^{-1}(t)\}) = F_X(F_X^{-1}(t)) = t \quad \text{for } 0 \leq t \leq 1$$

So $F_X(X)$ is a uniform random variable $U$ on $[0,1]$. Therefore, computing $X = F_X^{-1}(U)$ does what we want (and this can be generalized to the case where $F_X$ is not strictly increasing). In the present case, one can check that

$$F_X^{-1}(t) = \frac{1}{\lambda} \log \left( \frac{1}{1-t} \right) \quad \text{for } 0 \leq t \leq 1$$

so $X = \frac{1}{\lambda} \log \left( \frac{1}{U} \right)$ gives what we want (observing that $U$ and $1-U$ have the same distribution).

b) See the notebook. When $\lambda \leq 1$, the expectation of $Y$ is infinite (see part d) below), so the sequence $\hat{\mu}_n$ diverges as $n$ gets large. The situation is similar for $\hat{\sigma}_n$ when $\lambda \leq 2$ (and even worse when $\lambda \leq 1$).

c) See the notebook (and in particular the graph in there).

d) The cdf of $Y$ is given by

$$F_Y(t) = F_X(\log(t)) = 1 - \exp(-\lambda \log(t)) = 1 - \frac{1}{t^\lambda} \quad \text{if } t \geq 1 \quad \text{and} \quad F_Y(t) = 0 \quad \text{if } t \leq 1$$

and its pdf is given correspondingly by $p_Y(t) = \lambda / t^{\lambda+1}$ for $t \geq 1$ (and $p_Y(t) = 0$ for $t \leq 1$). This is the Pareto distribution with parameter $\lambda$.

In order to compute $\mathbb{E}(Y)$ and $\text{Var}(Y)$, it is easier to use the fact that $Y = e^X$:

$$\mathbb{E}(Y) = \mathbb{E}(e^X) = \int_0^\infty e^x \lambda e^{-\lambda x} \, dx = \frac{\lambda}{\lambda - 1}$$

for $\lambda > 1$ (and $\mathbb{E}(Y) = +\infty$ for $\lambda \leq 1$). Similarly, we compute

$$\mathbb{E}(Y^2) = \mathbb{E}(e^{2X}) = \int_0^\infty e^{2x} \lambda e^{-\lambda x} \, dx = \frac{\lambda}{\lambda - 2}$$

for $\lambda > 2$ (and $\mathbb{E}(Y) = +\infty$ for $\lambda \leq 2$). From this, we deduce that $\text{Var}(Y)$ is ill defined when $\lambda \leq 2$ and

$$\text{Var}(Y) = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 = \frac{\lambda}{\lambda - 2} - \left( \frac{\lambda}{\lambda - 1} \right)^2$$

$$= \lambda \left( \frac{(\lambda - 1)^2 - \lambda (\lambda - 2)}{(\lambda - 2) (\lambda - 1)^2} \right) = \frac{\lambda}{(\lambda - 2) (\lambda - 1)^2}$$

when $\lambda > 2$. 
Exercise 3. a) The computation of the characteristic function gives in this case:
\[ \phi_X(t) = \sum_{k \geq 0} \frac{\lambda^k e^{-\lambda}}{k!} e^{itk} = \sum_{k \geq 0} \frac{(\lambda e^{it})^k e^{-\lambda}}{k!} = e^{\lambda e^{it}} e^{-\lambda} = e^{\lambda(e^{it} - 1)} \]

b) The general expression for \( \phi_X \) is given by
\[ \phi_X(t) = \sum_{\ell \in \mathbb{Z}} \mathbb{P}(\{X = \ell\}) e^{i\ell t} \]

Plugging this expression into the proposed formula, we find
\[
\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-itk} \phi_X(t) \, dt = \sum_{\ell \in \mathbb{Z}} \mathbb{P}(\{X = \ell\}) \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-it(\ell - k)} \, dt = \sum_{\ell \in \mathbb{Z}} \mathbb{P}(\{X = \ell\}) \delta_{k\ell} = \mathbb{P}(\{X = k\})
\]
where we have switched the sum and integral without too much checking and we have used the fact that for \( k \neq \ell \):
\[
\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-it(\ell - k)} \, dt = \frac{1}{2\pi i(\ell - k)} \bigg|_{t = \pi}^{t = -\pi} = 0
\]

c) Let us compute
\[
\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-itk} \cos(t) \, dt = \frac{1}{4\pi} \int_{-\pi}^{\pi} e^{-itk} (e^{it} + e^{-it}) \, dt
\]
\[
= \frac{1}{4\pi} \int_{-\pi}^{\pi} (e^{-it(k-1)} + e^{-it(k+1)}) \, dt = \begin{cases} 1/2 & \text{if } k \in \{-1, +1\} \\ 0 & \text{otherwise} \end{cases}
\]
by the same argument as above.

BONUS d) We know that \( \phi_X(t) = \cos(t) \) is a characteristic function because \( \phi_X(0) = \cos(0) = 1 \), \( \phi_X \) is continuous on \( \mathbb{R} \), and also positive semi-definite. Indeed, using the trigonometric identity \( \cos(a - b) = \cos(a) \cos(b) + \sin(a) \sin(b) \), we obtain
\[
\sum_{j,k=1}^{n} c_j \overline{c_k} \phi_X(t_j - t_k) = \sum_{j,k=1}^{n} c_j \overline{c_k} \cos(t_j - t_k) = \sum_{j,k=1}^{n} c_j \overline{c_k} (\cos(t_j) \cos(t_k) + \sin(t_j) \sin(t_k))
\]
\[
= \left| \sum_{j=1}^{n} c_j \cos(t_j) \right|^2 + \left| \sum_{j=1}^{n} c_j \sin(t_j) \right|^2 \geq 0
\]
for every \( n \geq 1 \), \( t_1, \ldots, t_n \in \mathbb{R} \) and \( c_1, \ldots, c_n \in \mathbb{C} \).
Exercise 4. a) i) From the course, we know that if $E(|X|) < +\infty$, then $\phi_X$ is continuously differentiable on $\mathbb{R}$. Using the contraposition, we deduce that $E(|X|) = +\infty$ here.

a) ii) The fact that $\phi_X$ is integrable on $\mathbb{R}$ implies that $X$ admits a pdf $p_X$.

b) By the inversion formula seen in class, we have

$$p_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} e^{-\lambda|t|} dt = \frac{1}{2\pi} \left( \int_{-\infty}^{0} e^{-t(ix-\lambda)} dt + \int_{0}^{+\infty} e^{-t(ix+\lambda)} dt \right)$$

$$= \frac{1}{2\pi} \left( -\frac{1}{ix-\lambda} + \frac{1}{ix+\lambda} \right) = \frac{1}{\pi} \frac{\lambda}{x^2 + \lambda^2}$$

This pdf is the that of a (centered) Cauchy distribution with parameter $\lambda$ (also known as Lorentz distribution in physics). The word “centered” is a bit misleading here, as we have seen in part a)i) that $E(|X|) = +\infty$ (which can also been checked directly from the expression of $p_X$), so that $E(X)$ is ill-defined. Nevertheless, the pdf appears to have a peak clearly centered in $x = 0$ here, and writing $E(X) = 0$ can actually be justified via a more general definition of expectation. Besides, the parameter $\lambda > 0$ is connected to the width of the peak, but is by no means connected to the standard deviation of the random variable $X$, which is truly infinite.

c) Using the change of variable formula, we obtain

$$p_Y(x) = p_{1/X}(x) = p_X \left( \frac{1}{x} \right) \cdot \left| \frac{1}{x^2} \right| = \frac{1}{\pi} \frac{\lambda}{x^{-2} + \lambda^2} \frac{1}{x^2}$$

so we see that $Y$ is again a Cauchy random variable, with parameter $1/\lambda$.

d) By the factorization property of characteristic functions, we obtain

$$\phi_{X_1+\ldots+X_n}(t) = \prod_{i=1}^{n} \phi_{X_i}(t) = (\phi_{X}(t))^n = \exp(-\lambda n|t|)$$

so $X_1 + \ldots + X_n$ is also a Cauchy random variable with parameter $\lambda n$, and $Z_n = \frac{X_1+\ldots+X_n}{n}$ is a Cauchy random variable with parameter $\lambda$, for every $n \geq 1$. Similarly, we obtain, using part b),

$$\phi_{1/X_1+\ldots+1/X_n}(t) = (\phi_{1/X}(t))^n = \exp(-n|t|/\lambda)$$

so $1/X_1+\ldots+1/X_n$ is a Cauchy random variable with parameter $n/\lambda$. Therefore, again by part b), $1/X_1+\ldots+1/X_n$ is a Cauchy random variable with parameter $\lambda/n$ and $W_n = 1/X_1+\ldots+1/X_n$ is (again) a Cauchy random variable with parameter $\lambda$.

e) The first oddity of the above results is that the empirical average $Z_n = \frac{X_1+\ldots+X_n}{n}$ does not converge to a limit as $n$ goes to infinity. One reason for this is that $E(|X|) = +\infty$, so the law of large numbers does not hold, as we shall see later in the course. The second oddity is that the sum of an arbitrary number of Cauchy random variables is still a Cauchy random variable. The other well known distribution sharing this property is the Gaussian distribution, but that’s basically it, as this property is an exception among probability distributions. The third oddity is that the arithmetic mean $Z_n$ of the random variables $X_1, \ldots, X_n$ has the same distribution as their harmonic mean $W_n$. However, as we deal here with random variables taking positive and negative values, the classical inequality “arithmetic mean $\geq$ harmonic mean” does not hold, so there is no contradiction.