Statistics 2 Unit 1 Team 8

Nikolaos Kornilakis Rodrigo Viale Jakub Trnan Aleksandra Daneva
Luis Diego Pena Monge

2024-03-13

Exercise 1

Let $X(n,p) \sim Binom(n,p)$. Then $P(X=k) = \binom{n}{k} p^k q^{n-k}$. We have that:

$$\phi_{X(n,p)} = E[e^{itX}] = sum_{k=0}^{\infty} e^{itk} \binom{n}{k} p^k q^{n-k}$$

$$= \sum_{k=0}^{\infty} \binom{n}{k} (e^{it}p)^k q^{n-k}$$

Through the binomial theorem:

$$= (pe^{it} + q)^n$$

$$= (1 + p(e^{it} - 1))^n$$

Applying the limits as $n \to \infty$ and $p \to 0$ we get:

$$\lim_{n\to\infty, p\to 0} \phi_{X(n,p)} = \lim_{n\to\infty, p\to 0} (1 + p(e^{it} - 1))^n$$

$$= \lim_{n\to\infty, p\to 0} \left(1 + \frac{np}{n}(e^{it} - 1)\right)^n$$

Since $np \to \lambda$:

$$= e^{\lambda(e^{it}-1)}$$

As we showed in class, this last function is the characteristic function of a Poisson random variable with parameter λ. Note that this function is continuous everywhere, particularly at the origin. By Levy's continuity theorem, the sequence of binomials with parameters n, p converge in distribution to a Poisson with parameter λ when $n \to \infty$ and $p \to 0$, which is what we wanted to show.

Exercise 2

Let $X(\alpha) \sim Gamma(\alpha, \lambda)$ where α is the shape and λ the rate parameter respectively. Then we have that $f_{X(\alpha)}(x) = \frac{\lambda}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \mathbb{I}_{\{x>0\}}$. As we have previously shown in Statistics 1, $E[X(\alpha)] = \frac{\alpha}{\lambda}$ and $Var[X(\alpha)] = \frac{\alpha}{\lambda^2}$.

Now, the characteristic function of the **non-standardized** variable is given by:

$$
\begin{aligned}
\phi_{X(\alpha)}(t) = E[e^{itX(\alpha)}] &= \int_{\mathbb{R}} e^{itx}\frac{\lambda}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}\mathbb{I}_{\{x>0\}} \\
&= \int_0^{\infty} e^{itx}\frac{\lambda}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x} \\
&= \frac{\lambda^{\alpha}}{\Gamma(\alpha)}\int_0^{\infty} x^{\alpha-1}e^{-x(\lambda-it)}dx \\
&\text{Substituting } u = x(\lambda-it) \Rightarrow du = (\lambda-it)dx \\
&= \frac{\lambda^{\alpha}}{\Gamma(\alpha)}\int_0^{\infty}\left(\frac{u}{\lambda-it}\right)^{\alpha-1}e^{-u}\frac{1}{\lambda it}du \\
&= \frac{\lambda^{\alpha}}{\Gamma(\alpha)(\lambda-it)^{\alpha}}\int_0^{\infty} u^{\alpha-1}e^{-u}du \\
&= \frac{\lambda^{\alpha}}{\Gamma(\alpha)(\lambda-it)^{\alpha}}\Gamma(\alpha) \\
&= \left(\frac{\lambda}{\lambda-it}\right)^{\alpha} \\
&= \left(1-\frac{it}{\lambda}\right)^{-\alpha}
\end{aligned}
$$

Now let's standardize the random variable. Let $Z(\alpha) = \frac{X(\alpha)-\frac{\alpha}{\lambda}}{\frac{\sqrt{\alpha}}{\lambda}}$. Let $\mu(\alpha) = \frac{\alpha}{\lambda}$ and $\sigma(\alpha) = \frac{\sqrt{\alpha}}{\lambda}$. As seen in class:

$$
\begin{aligned}
\phi_{Z(\alpha)}(t) &= e^{-it\frac{\mu(\alpha)}{\sigma(\alpha)}}\phi_{X(\alpha)}\left(\frac{t}{\sigma(\alpha)}\right) \\
&= e^{-it\sqrt{\alpha}}\left(1-\frac{it}{\sqrt{\alpha}}\right)^{-\alpha}
\end{aligned}
$$

Now notice that:

$$
\begin{aligned}
\left(1-\frac{it}{\sqrt{\alpha}}\right)^{-\alpha} &= e^{-\alpha\ln\left(1-\frac{it}{\sqrt{\alpha}}\right)} \\
&\text{Using the Taylor expansion of } \ln(1-x): \\
&= e^{-\alpha\left[\frac{-it}{\sqrt{\alpha}}-\frac{1}{2}\left(\frac{it}{\sqrt{\alpha}}\right)^2+o\left(\left(\frac{it}{\sqrt{\alpha}}\right)^2\right)\right]} \\
&= e^{it\sqrt{\alpha}-\frac{1}{2}t^2+o(t^2)}
\end{aligned}
$$

Then, multiplying with the first term from above we get:

$$
\begin{aligned}
\phi_{Z(\alpha)}(t) &= e^{-it\sqrt{\alpha}}e^{it\sqrt{\alpha}-\frac{1}{2}t^2+o(t^2)} \\
&= e^{\frac{-1}{2}t^2+o(t^2)}
\end{aligned}
$$

Notice the second term is negligible in the limit, therefore:

$$\lim_{\alpha\to\infty} \phi_{Z(\alpha)}(t) = e^{\frac{-1}{2}t^2}$$

Notice that this is the characteristic function of a standard normal variable. All previous functions (particularly the limit) is continuous everywhere, and by Levy continuity theorem, the desired convergence is proven.

Exercise 3

Let R_i be the i-th roundup error. We examine

$$\sum_{i=1}^{100} R_i$$

where $R_i \sim U(-0.5, 0.5)$

First we calculate expectation:

$$\mathbb{E}\left[\sum_{i=1}^{100} R_i\right] = \sum_{i=1}^{100} \mathbb{E}\left[R_i\right] = \sum_{i=1}^{100} \int_{-0.5}^{0.5} x dx = \sum_{i=1}^{100} 0 = 0$$

Now, assuming independence (because we'll be using the CLT), we calculate variance:

$$\mathbb{V}\left[\sum_{i=1}^{100} R_i\right] = \sum_{i=1}^{100} \mathbb{V}\left[R_i\right]$$

$$\mathbb{V}\left[R_i\right] = \mathbb{E}\left[R_i^2\right] - \mathbb{E}\left[R_i\right]^2 = \int_{-0.5}^{0.5} x^2 dx - 0 = \frac{1}{12}$$

So our $\sigma = \frac{1}{\sqrt{12}}$. Now we need to standardize our round-off errors. Define

$$Z = \frac{\sum_{i=1}^{100} R_i - 100 \times 0}{\frac{1}{\sqrt{12}} \times \sqrt{100}} = \frac{\sqrt{3}}{5} \sum_{i=1}^{100} R_i$$

Now we know that by Central Limit Theorem, we can approximate \$Z \$ using a $N(0,1)$ random variable (as asymptotically this approximation holds). Let $x \geq 0$. Then:

$$\begin{aligned} P\left(|\sum_{i=1}^{100} R_i| > x\right) &= P\left(\sum_{i=1}^{100} R_i > x\right) + P\left(\sum_{i=1}^{100} R_i < -x\right) \\ &= P\left(Z > \frac{\sqrt{3}}{5} x\right) + P\left(Z < \frac{-\sqrt{3}}{5} x\right) \\ &= 1 - \Phi\left(\frac{\sqrt{3}}{5}\right) + \Phi\left(\frac{-\sqrt{3}}{5}\right) \\ &= 2\left(1 - \Phi\left(\frac{\sqrt{3}}{5}\right)\right) \end{aligned}$$

We calculate the respective results for $x = 1, 2, 5$:

```r
# x = 1
2*(1-pnorm(sqrt(3)/5))
```

```
## [1] 0.7290345
```

```r
# x = 2
2*(1-pnorm(2*(sqrt(3)/5)))
```

```
## [1] 0.4884223
```

```r
# x = 5
2*(1-pnorm(5*(sqrt(3)/5)))
```

```
## [1] 0.08326452
```

Exercise 4

To approximate $P(S \leq 10)$ using CLT, where $X_1, ..., X_{20}$ are i.i.d. with density function $f(x) = 2x$ for $0 \leq x \leq 1$ and $S = X_1 + ... + X_{20}$, let's first find the mean and variance of S.

Note

$$E[X_i] = \int_0^1 2x^2 ds = 2\left[\frac{x^3}{3}\right]_0^1 = \frac{2}{3}$$

$$E[X_i^2] = \int_0^1 2x^3 ds = 2\left[\frac{x^4}{4}\right]_0^1 = \frac{1}{2}$$

Then

$$Var[X_i] = \frac{1}{2} - \frac{4}{9} = \frac{1}{18} \Rightarrow \sigma = \frac{\sqrt{2}}{6}$$

By CLT we can approximate $\frac{S-20\frac{2}{3}}{\sqrt{20}\frac{\sqrt{2}}{6}}$ by a standard normal distribution. Notice $\frac{10-20\frac{2}{3}}{\sqrt{20}\frac{\sqrt{2}}{6}} = -\sqrt{10}$. Therefore:

$$\begin{aligned} P(S \leq 10) &= P\left(\frac{S - \frac{40}{3}}{\frac{\sqrt{10}}{3}} \leq -\sqrt{10}\right) \\ &\approx \Phi(-\sqrt{10}) \end{aligned}$$

The value of the probability would be:

```r
pnorm(-sqrt(10))
```

```
## [1] 0.0007827011
```

Exercise 5

a)

Let's first calculate the real value of the integral:

Let $u = 2\pi x$. Then:

$$
\begin{aligned}
\int_0^1 \cos(2\pi x)dx &= int_0^{2\pi} \frac{\cos(u)}{2\pi} du \\
&= \frac{1}{2\pi} \sin(u)\Big|_0^{2\pi} \\
&= 0
\end{aligned}
$$

Notice that for the Monte Carlo estimation, we can draw from a standard uniform distribution, i.e., $f(x) = \mathbb{I}_{[0,1]}$ and $g(x) = \cos(2\pi x)$.

```r
set.seed(8888)

n <- 100
mc_100 <- mean(cos(2*pi*runif(n)))

n <- 1000
mc_1000 <- mean(cos(2*pi*runif(n)))

real <- 0

data.frame(n = c("Real", "100", "1000"),
           Value = c(real, mc_100, mc_1000),
           Abs.Diff = abs(c(real - real, mc_100 - real, mc_1000 - real)))
```

```
##      n       Value    Abs.Diff
## 1 Real 0.000000000 0.000000000
## 2  100 0.054222406 0.054222406
## 3 1000 0.001796749 0.001796749
```

Clearly there is an improvement by increasing the value of n, which follows from the fact that the monte carlo estimation converges to the real value as $n \to \infty$.

b)

For the case of $\int_0^1 cos(2\pi x^2)dx$, it is important to note that this integral does not have a closed form solution. Therefore, we take as real value the numerical approximation provided by WolframAlfa up to 10 decimal points:

```r
set.seed(8888)
# WolframAlfa value
real <- 0.2441267030

n <- 100
mc_100 <- mean(cos(2*pi*runif(n)^2))

n <- 1000
mc_1000 <- mean(cos(2*pi*runif(n)^2))

data.frame(n = c("Real", "100", "1000"),
           Value = c(real, mc_100, mc_1000),
           Abs.Diff = abs(c(real - real, mc_100 - real, mc_1000 - real)))
```

```
##      n     Value    Abs.Diff
```

```
## 1 Real 0.2441267 0.000000000
## 2  100 0.3174299 0.073303229
## 3 1000 0.2412027 0.002924026
```

Again, the value obtained by increasing the number of draws from the standard uniform r.v. is a better approximation.

Exercise 6

$$\begin{aligned}
\text{Var}(\hat{\theta}) &= \text{Var}\left(\frac{1}{n}\sum_{i=1}^{n} g(x_i)\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} \text{Var}(g(x_i)) \\
&= \frac{1}{n}\text{Var}(g(x_i))
\end{aligned}$$

Implementing this in R for $\theta = \int_0^1 cos(2\pi x)dx$:

```r
set.seed(8888)
ns <- c(100, 1000, 10000, 10000)

res <- lapply(ns, function(n){
  est <- cos(2 * pi * runif(n))
  mc <- mean(est)
  mc_var <- var(est)/n
  abs_error <- abs(0 - mc)

  return(data.frame(n = n,
                    MC = mc,
                    Std.Err = sqrt(mc_var),
                    Abs.Err = abs_error))
})

do.call(rbind, res)
```

```
##       n           MC     Std.Err     Abs.Err
## 1   100  0.054222406 0.072474888 0.054222406
## 2  1000  0.001796749 0.022512409 0.001796749
## 3 10000 -0.005075061 0.007044437 0.005075061
## 4 10000  0.001183020 0.007061683 0.001183020
```

We see that the absolute errors in our realizations are lower (sometimes considerably) than the standard error of the monte carlo approximation.

Exercise 7

a)

$$\begin{aligned}
\mathbb{E}[\hat{\theta}] &= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{g(x_i)}{f(x_i)}\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\frac{g(x_i)}{f(x_i)}\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_a^b \frac{g(x)}{f(x)} f(x)\,dx_i \\
&= \frac{1}{n}\cdot n \cdot \int_a^b g(x)\,dx \\
&= \theta
\end{aligned}$$

b)

$$\begin{aligned}
\mathrm{Var}(\hat{\theta}) &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{g(x_i)}{f(x_i)}\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}\left(\frac{g(x_i)}{f(x_i)}\right) \\
&= \frac{1}{n}\mathrm{Var}\left(\frac{g(x)}{f(x)}\right)
\end{aligned}$$

Now, note that

$$E\left[\left(\frac{g(x)}{f(x)}\right)^2\right] = \int_a^b \frac{g^2(x)}{f^2(x)} f(x)dx = \int_a^b \frac{g^2(x)}{f(x)}dx$$

Then:

$$\mathrm{Var}\left(\frac{g(x)}{f(x)}\right) = \int_a^b \frac{g^2(x)}{f(x)}\,dx - \theta^2$$

And therefore:

$$\mathrm{Var}(\hat{\theta}) = \frac{1}{n}\left(\int_a^b \frac{g^2(x)}{f(x)}\,dx - \theta^2\right)$$

Now, let's see an example of finite variance. Take $f(x) = \frac{1}{b-a}\mathbb{I}_{[a,b]}$. We want to find g such that

$$\left|\int_a^b g(x)dx\right| < \infty \text{ and } \left|\int_a^b g^2(x)dx\right|$$

Consider $g(x) = a_0 + a_1 x + \ldots + a_m x^m$ any non-constant real-coefficient, m-degree polynomial. Then $g^2(x)$ is a degree $2m$ polynomial. Notice that g and g^2 are continuous (and therefore bounded) on $[a, b]$, so both their integrals converge. Therefore,

$$\mathrm{Var}(\hat{\theta}) = \frac{1}{n}\left(\int_a^b \frac{g^2(x)}{f(x)}\,dx - \left(\int_a^b g(x)\,dx\right)^2\right)$$

is finite.

For the infinite case, again take $f(x) = \frac{1}{b-a}\mathbb{I}_{[a,b]}$. Here, we want to find a function g such that

$$\left| \int_a^b g(x)dx \right| < \infty \text{ but } \int_a^b g^2(x)dx \text{ diverges}$$

Since we want the second integral to diverge, consider the case $a = 0, b = 1$ and take $g^2(x) = \frac{1}{x}$. Then:

$$\int_0^1 g^2(x)dx = \int_0^1 \frac{1}{x}dx = \ln(x)|_0^1 = +\infty$$

However, $\int_0^1 g(x)dx = \int_0^1 \frac{1}{\sqrt{x}}dx = 2\sqrt{x}|_0^1 = 2\sqrt{2} < \infty$. In this case:

$$\text{Var}(\hat{\theta}) = \frac{1}{n}\left(\int_a^b \frac{g^2(x)}{f(x)}\, dx - \left(\int_a^b g(x)\, dx \right)^2 \right) = +\infty$$

c)

We can improve this estimate by choosing a probability density function that has a similar shape to $g(x)$, such as an exponential distribution restricted to the interval $[0, 1]$. By choosing $f(x)$ to resemble $g(x)$, we ensure that more samples are taken where the integrand is larger, which can lead to a more accurate estimate for the same number of samples. Below, we show this method choosing a an exponential distribution with $\lambda = 1$ as $f(x)$, and comparing it with a the same estimate using a standard uniform distribution.

To draw from the exponential distribution, notice that $f(x) = e^{-x}$, and we must then have $g(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\frac{1}{e^{-x}}$ if x in $[0, 1]$, 0 otherwise.

Another option is to take advantage of the fact that the given function is already a density, that of a standard normal.

```r
set.seed(8888)
n <- 1000
real <- pnorm(1) - pnorm(0)

# Uniform
mc_unif <- mean(1/sqrt(2*pi) * exp(-runif(n)^2/2))

# Exponential
x <- rexp(n, 1)
g <- (x <= 1) * 1/sqrt(2*pi) * exp(-x^2/2) * 1/exp(-x)
mc_exp <- mean(g)

# Normal
x <- rnorm(n)
mc_norm <- mean(x >= 0 & x <= 1)

# Error measures
data.frame(Distribution = c("Real value", "Uniform", "Exponential", "Normal"),
           Estimate = c(real, mc_unif, mc_exp, mc_norm),
           Abs.Error = abs( c(0, real - mc_unif, real - mc_exp, real - mc_norm)),
           Rel.Error = abs( c(0, real - mc_unif, real - mc_exp, real - mc_norm))/real)
```

```
##    Distribution  Estimate    Abs.Error   Rel.Error
## 1   Real value 0.3413447 0.0000000000 0.000000000
## 2      Uniform 0.3423292 0.0009844975 0.002884173
## 3  Exponential 0.3560226 0.0146778955 0.043000209
## 4       Normal 0.3330000 0.0083447461 0.024446681
```

As we can see, using a uniform distribution seems to work best given the proposed alternatives. This can be explained through the fact that (as seen in class last semester), by sampling from distributions that take values outside of the range of interest and filtering only those observations that fall in this range, we might be throwing away additional points that would otherwise add information to our estimate and make it more precise. In other words, sampling n observations of a distribution that takes values outside of $[0, 1]$ does not mean that we have n useful observations providing information to the approximation. While there might be particular scenarios in which a single sample performs excellently, this is by no means the norm.

Exercise 8

Monte Carlo estimator is

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} cos(2\pi x_i)$$

where $X_i \sim U(0, 1)$ We compute the expected value of our random variable:

$$\mathbb{E}\left[cos(2\pi x_i)\right] = \int_0^1 cos(2\pi x_i)dx = 0$$

$$\mathbb{V}\left[cos(2\pi x_i)\right] = \int_0^1 cos^2(2\pi x_i)dx - 0$$

Substitute: $u = 2\pi x$ $du = 2\pi dx$

$$= \int_0^1 \frac{cos^2(u)}{2\pi}du = \frac{1}{2\pi}\int_0^1 cos^2(u) = \frac{1}{4\pi}\left[u + sin(u)cos(u)\right]_0^{2\pi} = \frac{2\pi}{4\pi} = \frac{1}{2}$$

Now by CLT:

$$\frac{\frac{1}{n}\sum_{i=1}^{n} cos(2\pi x_i) - \mathbb{E}\left[cos(2\pi x_i)\right]}{\sqrt{\frac{1}{n}\mathbb{V}\left[cos(2\pi x_i)\right]}} \sim N(0, 1)$$

This can be written down as:

$$\frac{\hat{\theta} - \theta}{\frac{1}{\sqrt{2000}}} \sim N(0, 1)$$

Now, $|\hat{\theta} - \theta| \le \Delta$, so $-\Delta \le \hat{\theta} - \theta$ and $\hat{\theta} - \theta \le \Delta$ So, $\mathbb{P}\left(|\hat{\theta} - \theta| \le \Delta\right)$

$$= \mathbb{P}\left(-\Delta \le \hat{\theta} - \theta \le \Delta\right) = \mathbb{P}\left(-\Delta\sqrt{2000} \le \sqrt{2000}(\hat{\theta} - \theta) \le \Delta\sqrt{2000}\right)$$

$$= \phi(\Delta\sqrt{2000}) - \phi(-\Delta\sqrt{2000})$$

$$= 2\phi(\Delta\sqrt{2000}) - 1 = 0.05$$

$$\frac{\phi^{-1}(\frac{1.05}{2})}{\sqrt{2000}} = \Delta$$

```r
qnorm(1.05/2) / sqrt(2000)
```

```
## [1] 0.001402166
```

Exercise 9

Let $U_1, U_2, \ldots, U_n$ be i.i.d. $U(0,1)$ and $U_{(n)} = \max\{U_1, \ldots, U_n\}$.

The cumulative distribution function $F_{U_{(n)}}(x)$ is given by:

$$\begin{aligned}
F_{U_{(n)}}(x) &= P(U_{(n)} \le x) \\
&= P(\max\{U_1, \ldots, U_n\} \le x) \\
&= P(U_1 \le x, \ldots, U_n \le x) \\
&= \prod_{i=1}^{n} P(U_i \le x) \\
&= \begin{cases} 1 & x \ge 1 \\ x^n & x \in [0,1] \\ 0 & x < 0 \end{cases}
\end{aligned}$$

The probability density function $f_{U_{(n)}}(x)$ is the derivative of $F_{U_{(n)}}(x)$:

$$\begin{aligned}
f_{U_{(n)}}(x) &= \frac{d}{dx} F_{U_{(n)}}(x) \\
&= \begin{cases} nx^{n-1} & x \in [0,1] \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

The expected value $E[U_{(n)}]$:

$$\begin{aligned}
E[U_{(n)}] &= \int_0^1 nx^n dx \\
&= n \left[\frac{x^{n+1}}{n+1} \right]_0^1 \\
&= \frac{n}{n+1}
\end{aligned}$$

The second moment $E[U_{(n)}^2]$:

$$\begin{aligned}
E[U_{(n)}^2] &= \int_0^1 nx^{n+1} dx \\
&= n \left[\frac{x^{n+2}}{n+2} \right]_0^1 \\
&= \frac{n}{n+2}
\end{aligned}$$

The variance $\text{Var}(U_{(n)})$:

$$\begin{aligned}
\text{Var}(U_{(n)}) &= E[U_{(n)}^2] - (E[U_{(n)}])^2 \\
&= \frac{n}{n+2} - \left(\frac{n}{n+1} \right)^2 \\
&= \frac{n}{(n+2)(n+1)^2}
\end{aligned}$$

Standardizing, we get $Z_{(n)}$:

$$\begin{aligned}
Z_{(n)} &= \frac{U_{(n)} - E[U_{(n)}]}{\sqrt{\mathrm{Var}(U_{(n)})}} \\
&= \frac{U_{(n)} - \frac{n}{n+1}}{\sqrt{\frac{n}{(n+2)(n+1)^2}}} \\
&= \frac{\sqrt{(n+2)(n+1)^2}}{\sqrt{n}} U_{(n)} - \frac{n\sqrt{(n+2)(n+1)^2}}{(n+1)\sqrt{n}} \\
&= \frac{(n+1)\sqrt{(n+2)}}{\sqrt{n}} U_{(n)} - \sqrt{n(n+2)}
\end{aligned}$$

Thus, the probability $P(Z_{(n)} \leq x)$ is:

$$\begin{aligned}
P(Z_{(n)} \leq x) &= P\left(\frac{(n+1)\sqrt{n+2}U_{(n)} - \sqrt{n(n+2)}}{\sqrt{n}} \leq x\right) \\
&= P\left(U_{(n)} \leq \sqrt{n}\left(x + \sqrt{\frac{n(n+2)}{(n+1)^2}}\right)\right) \\
&= P\left(U_{(n)} \leq \sqrt{n}\frac{x}{(n+1)\sqrt{n+2}} + \frac{n}{n+1}\right) \\
&= \left(\frac{x\sqrt{n}}{(n+1)\sqrt{n+2}} + \frac{n}{n+1}\right)^n \\
&= \left(\frac{n}{n+1}\right)^n \left(\frac{x}{\sqrt{n(n+2)}} + 1\right)^n \\
&= \left(1 + \frac{1}{n}\right)^{-n} \left(\frac{x}{\sqrt{n(n+2)}} + 1\right)^n \\
&\text{Taking the limit as } n \to \infty \\
&\to e^{-1} \cdot e^x = e^{x-1}
\end{aligned}$$

Exercise 11

a)

$$\begin{aligned}
K_1 &= \frac{d}{dt}k(t)\Big| t = c = \frac{d}{dt}\log(m(t))\Big|_{t=0} \\
&= \frac{1}{m(t)} \cdot \frac{dm(t)}{dt}\Big|_{t=0} \\
&= \frac{1}{m(0)} \cdot E[X] = E[X]
\end{aligned}$$

$$\begin{aligned}
K_2 &= \frac{d^2}{dt^2}k(t)\Big|_{t=0} = \frac{d}{dt}\left(\frac{1}{m(t)} \cdot \frac{dm(t)}{dt}\right)\Big|_{t=0} \\
&= -\frac{dm(t)}{dt}\frac{1}{\mu(t)^2} \cdot \frac{dm(t)}{dt} + \frac{1}{m(t)} \cdot \frac{d^2m(t)}{dt^2}\Big|_{t=0} \\
&= -E[X]^2 + E[X^2] = E[X^2] - E[X]^2
\end{aligned}$$

$$\begin{aligned}
K_3 &= \frac{d^3}{dt^3}k(t)\Big|_{t=0} = \frac{d}{dt}\left(-\frac{dm(t)}{dt}\frac{1}{m(t)^2} \cdot \frac{dm(t)}{dt} + \frac{1}{m(t)} \cdot \frac{d^2m(t)}{dt^2}\right)\Big|_{t=0} \\
&= -2m'(t)m''(t)m(t) + 2m'(t)^3 + m(t)m'''(t) - m''(t)m'(t)\Big|_{t=0} \\
&= -2E[X]E[X^2] + 2E[X]^3 + E[X^3] - E[X]E[X^2] \\
&= E[X^3] - 3E[X]E[X^2] + 2E[X]^3
\end{aligned}$$

$$\begin{aligned}
K_4 &= \frac{d}{dt}\left[-2m'(t)m'''(t)m(t)^2 + 2m'(t)^3 + \frac{m(t)m'''(t) - m''(t)m'(t)m'(t)m''(t)}{m(t)^2}\right]\Bigg|_{t=0} \\
&= -2\left(m(t)^3m'''(t) + 4m'(t)^4 + m(t)m'''(t)m'(t) - m(t)\left(2m(t)m'(t)^3m''(t)\right)\right)/m(t)^5 + ... \\
&\quad ... + \left(m'''(t)m(t)^2 - m(t)m''(t)m''(t) - 2m'''(t)m(t)m'(t) + 2m'(t)^2m''(t)\right)/m(t)^3\Bigg|_{t=0} \\
&= E[X^4] - 4E[X^3]E[X] - 3E[X^2]^2 + 12E[X^2]E[X]^2 - 6E[X]^4
\end{aligned}$$

b)

For the first two, it is easy to notice that: $K_1 = E[X] = \mu_1$ and $K_2 = E[X^2] - E[X]^2 = \mu_2$. For the other two:

$$\begin{aligned}
\mu_3 &= E[(X - E[X])^3] \\
&= E[X^3 - 3X^2E[X] + 3XE[X]^2 - E[X]^3] \\
&= E[X^3] - 3E[X]E[X^2] + 2E[X]^3 \\
&= K_3
\end{aligned}$$

And finally, notice that:

$$
\begin{aligned}
E[(X-E[X])^4] &= E\left[X^4 - 4X^3E[X] + 6X^2E[X]^2 - 4XE[X]^3 + E[X]^4\right] \\
&= E[X^4] - 4E[X]E[X^3] + 6E[X^2]E[X]^2 - 4E[X]E[X^3] + E[X]^4
\end{aligned}
$$

$$
\begin{aligned}
3E\left[(X-E[X])^2\right]^2 &= 3\left(E[X^2] - E[X]^2\right)^2 \\
&= 3E[X^2]^2 - 6E[X^2]E[X]^2 + 3E[X]^4
\end{aligned}
$$

Taking the difference of the last two terms we get:

$$
\begin{aligned}
\mu_4 - 3\mu_3^2 &= E[X^4] - 4E[X]E[X^3] + 12E[X^2]E[X]^2 - 6E[X]^4 - 3E[X^2]^2 \\
&= K_4
\end{aligned}
$$

c)

Using the equivalences found in part b):

$$
\text{Skew} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{K_3}{K_2^{3/2}},
$$

$$
\text{Kurt} = \frac{\mu_4}{\mu_2^2} = \frac{K_4 + 3K_2^2}{K_2^2} = \frac{K_4}{K_2^2} + 3.
$$

Exercise 12

Poisson

Recall for a Poisson distribution the MGF is given by $m(t) = e^{\lambda(e^t-1)}$. Therefore:

$$
\begin{aligned}
E[X] &= m'(0) \\
&= e^{\lambda(e^t-1)}\lambda e^t|_{t=0} \\
&= \lambda
\end{aligned}
$$

$$
\begin{aligned}
E[X^2] &= m''(0) \\
&= e^{\lambda(e^t-1)}(\lambda e^t)^2 + e^{\lambda(e^t-1)}\lambda e^t|_{t=0} \\
&= \lambda^2 + \lambda
\end{aligned}
$$

$$
\begin{aligned}
E[X^3] &= m^{(3)}(0) \\
&= e^{\lambda(e^t-1)}\lambda e^t(\lambda e^t)^2 + 2(\lambda e^t)^2 e^{\lambda(e^t-1)} + e^{\lambda(e^t-1)}(\lambda e^t)^2 + \lambda e^t e^{\lambda(e^t-1)}|_{t=0} \\
&= \lambda^2 + 3\lambda^2 + \lambda
\end{aligned}
$$

$$
\begin{aligned}
E[X^4] &= m^{(4)}(0) \\
&= \lambda e^{\lambda(e^t-1)+t}(\lambda^3 e^{3t} + 6\lambda^2 e^{2t} + 7\lambda e^t + 1)|_{t=0} \\
&= \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda
\end{aligned}
$$

Therefore, substituting the moments in the formulas found in exercise 11 we get:

$$
\begin{aligned}
K_1 &= \lambda \\
K_2 &= \lambda^2 + \lambda - \lambda^2 = \lambda \\
K_3 &= \lambda^3 + 3\lambda^2 + \lambda - 3\lambda(\lambda^2 + \lambda) + 2\lambda^3 = \lambda \\
K_4 &= \lambda
\end{aligned}
$$

Normal

Recall for a normal distribution the MGF is given by $m(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$. Therefore:

$$
\begin{aligned}
E[X] &= m'(0) \\
&= e^{\mu t + \frac{\sigma^2 t^2}{2}}(\mu + \sigma^2 t)|_{t=0} \\
&= \mu
\end{aligned}
$$

$$
\begin{aligned}
E[X^2] &= m''(0) \\
&= e^{\mu t + \frac{\sigma^2 t^2}{2}}(\mu^2 + \sigma^2 + \sigma^4 t^2 + 2\mu\sigma^2 t)|_{t=0} \\
&= \mu^2 + \sigma^2
\end{aligned}
$$

$$
\begin{aligned}
E[X^3] &= m^{(3)}(0) \\
&= e^{\mu t + \frac{\sigma^2 t^2}{2}}(\mu + \sigma^2 t)(3\sigma^2 + (\mu + \sigma^2 t)^2)|_{t=0} \\
&= 3\mu\sigma^2 + \mu^3
\end{aligned}
$$

$$
\begin{aligned}
E[X^4] &= m^{(4)}(0) \\
&= e^{\mu t + \frac{\sigma^2 t^2}{2}}(3\sigma^4 + 6\sigma^2(\mu + \sigma^2 t)^2 + (\mu + \sigma^2 t)^4)|_{t=0} \\
&= 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4
\end{aligned}
$$

Once again, substituting the moments in the formulas found in exercise 11 we get:

$$
\begin{aligned}
K_1 &= \mu \\
K_2 &= \sigma^2 \\
K_3 &= 0 \\
K_4 &= 0
\end{aligned}
$$

Exercise 15

$$
\begin{aligned}
F(x) &= e^{-e^{-x}} \\
&\Rightarrow f(x) = \frac{d}{dx}F(x) = e^{-e^{-x}} \cdot e^{-x} \\
\text{Mode:} & \\
\frac{d}{dx}f(x) &= e^{-e^{-x}} \cdot (-e^{-x}) \cdot (e^{-x} - 1) = 0 \\
&\Rightarrow e^{-x} = 1 \\
&\Rightarrow x = 0 \\
\text{Therefore, mode} &= 0 \\
\text{Median:} & \\
F(x) &= e^{-e^{-x}} = \frac{1}{2} \\
&\Rightarrow -e^{-x} = \log\left(\frac{1}{2}\right) \\
&\Rightarrow e^{-x} = \log(2) \\
&\Rightarrow -x = \log(\log(2)) \\
&\Rightarrow x = -\log(\log(2))
\end{aligned}
$$

$$
\begin{aligned}
\text{MGF: } m(t) = \mathbb{E}[e^{tx}] &= \int_{-\infty}^{\infty} e^{tx} \cdot e^{-e^{-x}} \cdot e^{-x}\, dx \\
\text{Let } u &= e^{-x}, du = -e^{-x}dx \\
&= \int_{\infty}^{0} u^{-t} e^{-u}(-du) \\
&= \int_{0}^{\infty} u^{(1-t-1)} e^{-u}\, du \\
&= \Gamma(1-t)
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}[x] &= \frac{d}{dt}m(t)\Big|_{t=0} = -\Gamma'(1-t)\Big|_{t=0} = -\Gamma'(1) \\
\text{Now, } \Psi(z) &= \frac{\Gamma'(z)}{\Gamma(z)} \\
\Rightarrow \Gamma'(1) &= \Gamma(1) \cdot \Psi(1) = 0 \cdot (-\gamma) = \gamma \\
\text{Similarly, } \mathbb{E}[x^2] &= \frac{d^2}{dt^2}m(t)\Big|_{t=0} = -\frac{d^2}{dt^2}\Gamma(1-t)\Big|_{t=0} = -\Gamma''(1) \\
\Rightarrow \Gamma''(1-t) &= \Gamma'(1-t) \cdot \Psi(1-t) \\
\Rightarrow \Gamma''(1) &= -\Gamma'(1) \cdot \Psi'(1-t) \\
&= \Gamma(1-t)\Psi(1-t)\Psi'(1-t) - \Gamma(1-t)\Psi''(1-t)
\end{aligned}
$$

$$\Rightarrow \frac{d^2}{dt^2}m(t) = -\frac{d}{dt}\Gamma'(1-t) = \Gamma(1-t)\Psi(1-t) + \Gamma'(1-t)\Psi'(1-t)$$

$$\begin{aligned}
\text{Evaluating at } t=0 \text{ we get } \mathbb{E}[x^2] &= \Gamma(1)\Psi(1) + \Gamma(1)\Psi'(1) \\
&= \gamma^2 + \frac{\pi^2}{6} \\
\Rightarrow \text{Var}(x) &= \mathbb{E}[x^2] - (\mathbb{E}[x])^2 = \gamma^2 + \frac{\pi^2}{6} - \gamma^2 = \frac{\pi^2}{6} \\
\text{Now, } \mathbb{E}[e^{-x}] &= \Gamma(1-(-1)) = \Gamma(2) = 1 \cdot \Gamma(1) = 1
\end{aligned}$$

$$\begin{aligned}
\text{Now, notice that } \Gamma(z+1) &= z\Gamma(z) \\
&\Rightarrow \log(\Gamma(z+1)) = \log(z) + \log(\Gamma(z)) \\
&\Rightarrow \frac{d}{dz}\log(\Gamma(z+1)) = \frac{1}{z} + \frac{d}{dz}\log(\Gamma(z)) \\
&\Rightarrow \Psi(z+1) = \frac{1}{z} + \Psi(z) \\
&\Rightarrow \Psi(1) = \frac{1}{1} + \Psi(1) = 1 - \gamma \\
&\Rightarrow \Psi'(2) = -\frac{1}{1^2} + \Psi'(2) \\
\Rightarrow \mathbb{E}[xe^{-x}] &= \gamma - 1
\end{aligned}$$

$$\begin{aligned}
\text{Similarly: } \mathbb{E}[x^2e^{-x}] &= m''(-1) \\
&= \Gamma(2)\Psi(2)^2 + \Gamma(2)\Psi'(2) \\
&= 1 \cdot \Psi(1)^2 + 1 \cdot \left(-\frac{\pi^2}{6}\right) \\
&= (1-\gamma)^2 + \left(-\frac{\pi^2}{6}\right) \\
&= 1 - 2\gamma + \gamma^2 - \frac{\pi^2}{6} \\
&= \frac{\pi^2}{6} - 2\gamma + \gamma^2
\end{aligned}$$

Exercise 16

$$
\begin{aligned}
F(t) &= F_0\left(\frac{t-\mu}{\sigma}\right) \\
&= e^{-e^{-\left(\frac{t-\mu}{\sigma}\right)}} = \rho \\
\Leftrightarrow \quad & e^{-e^{-\left(\frac{t-\mu}{\sigma}\right)}} = \ln(\rho) \\
&\Leftrightarrow e^{-\left(\frac{t-\mu}{\sigma}\right)} = \ln\left(\frac{1}{\rho}\right) \\
&\Leftrightarrow -\left(\frac{t-\mu}{\sigma}\right) = \ln\left(\ln\left(\frac{1}{\rho}\right)\right) \\
&\Leftrightarrow t = -\sigma \ln\left(\ln\left(\frac{1}{\rho}\right)\right) + \mu
\end{aligned}
$$

Exercise 17

First of all, notice that:

$$
\begin{aligned}
F(x|\alpha,\lambda) &= \int_0^x \frac{\alpha\lambda^\alpha}{(\lambda+t)^{\alpha+1}}dt \\
&\text{Substituting } u = \lambda + t \\
&= \alpha\lambda^\alpha \int_\lambda^{\lambda+x} \frac{1}{u^{\alpha+1}}du \\
&= \alpha\lambda^\alpha \left[\frac{-u^{-\alpha}}{\alpha}\right]_\lambda^{\lambda+x} \\
&= 1 - \left(\frac{\lambda}{\lambda+x}\right)^\alpha
\end{aligned}
$$

Now, for the expected value:

$$
\begin{aligned}
E[X] &= \int_0^\infty x\frac{\alpha\lambda^\alpha}{(\lambda+x)^{\alpha+1}}dx \\
&= \alpha\lambda^\alpha \int_0^\infty \frac{x}{(\lambda+x)^{\alpha+1}}dx \\
&\text{Substituting } u = \lambda + x \\
&= \alpha\lambda^\alpha \left(\int_\lambda^\infty u^{-\alpha}du - \lambda\int_\lambda^\infty u^{-(\alpha+1)}du\right) \\
&= \alpha\lambda^\alpha \left(\frac{\lambda^{1-\alpha}}{\alpha-1} - \lambda\frac{\lambda^{-\alpha}}{\alpha}\right) \\
&= \frac{\lambda}{\alpha-1}
\end{aligned}
$$

For the second moment:

$$
\begin{aligned}
E[X^2] = \alpha\lambda^\alpha \int_0^\infty & \frac{x^2}{(\lambda+x)^{\alpha+1}} dx \\
& \text{Substituting and expanding the remaining binom } u = \lambda + x \\
& = \alpha\lambda^\alpha \left(\int_\lambda^\infty u^{-(\alpha-1)} du - 2\lambda \int_\lambda^\infty u^{-\alpha} du + \lambda^2 \int_\lambda^\infty u^{-(\alpha+1)} du \right) \\
& = \alpha\lambda^\alpha \left(\frac{\lambda^{2-\alpha}}{\alpha-2} - 2\lambda \frac{\lambda^{1-\alpha}}{\alpha-1} \lambda^2 \frac{\lambda^{-\alpha}}{\alpha} \right) \\
& = \frac{2\lambda^2}{(\alpha-2)(\alpha-1)}
\end{aligned}
$$

This means the variance is given by:

$$
Var[X] = E[X^2] - E[X]^2 = \frac{2\lambda^2}{(\alpha-2)(\alpha-1)} - \left(\frac{\lambda}{\alpha-1} \right)^2 = \frac{\lambda^2\alpha}{(\alpha-2)(\alpha-1)^2}
$$

Finally, fot the excess, let $M \geq 0$. Recall:

$$
\begin{aligned}
P(X - M \leq x | X > M) &= P(X \leq x + M | X > M) \\
&= \frac{P(M < X \leq x + M)}{P(X > M)} \\
&= \frac{\int_M^{x+M} f(t)dt}{\int_M^\infty f(t)dt}
\end{aligned}
$$

Now:

$$
\int_M^{x+M} f(t)dt = \alpha\lambda^\alpha \int_M^{x+M} (\lambda+t)^{-(\alpha+1)} dt = \alpha\lambda^\alpha \left[\frac{-(\lambda+t)^{-\alpha}}{\alpha} \right]_M^{x+M} = \lambda^\alpha[(\lambda+M)^{-\alpha} - (\lambda+x+M)^{-\alpha}]
$$

And:

$$
\int_M^\infty f(t)dt = \alpha\lambda^\alpha \int_M^\infty (\lambda+t)^{-(\alpha+1)} dt = \alpha\lambda^\alpha \left[\frac{-(\lambda+t)^{-\alpha}}{\alpha} \right]_M^\infty = \lambda^\alpha(\lambda+M)^{-\alpha}
$$

Therefore:

$$
P(X - M \leq x | X > M) = \frac{\lambda^\alpha[(\lambda+M)^{-\alpha} - (\lambda+x+M)^{-\alpha}]}{\lambda^\alpha(\lambda+M)^{-\alpha}} = 1 - \frac{(\lambda+M+x)^{-\alpha}}{(\lambda+M)^{-\alpha}} = 1 - \left(\frac{\lambda+M}{\lambda+M+x} \right)^\alpha
$$

Which is exactly the distribution of a Lomax random variable with parameters α and $\lambda + M$.