STAT 511 su 2020 hw 6

mgfs, quantiles

1. Let $X \sim \text{Poisson}(\lambda)$.

   (a) Show that the mgf of X is given by $M_X(t) = e^{\lambda(e^t - 1)}$.

   We have

$$\begin{aligned}
{}_X(t) &= \mathbb{E}e^{tX} \\
&= \sum_{i=0}^{\infty} e^{tx} \cdot \frac{e^{-\lambda}\lambda^x}{x!} \\
&= \sum_{i=0}^{\infty} \frac{e^{-\lambda}(\lambda e^t)^x}{x!} \\
&= e^{\lambda e^t} e^{-\lambda} \underbrace{\sum_{i=0}^{\infty} \frac{e^{-\lambda e^t}(\lambda e^t)^x}{x!}}_{\text{sum over Poisson}(\lambda e^t)\text{ pmf}} \\
&= e^{\lambda(e^t - 1)}.
\end{aligned}$$

   (b) Use the mgf to find

   i. $\mathbb{E}X$.

   We have

$$M_X^{(1)}(t) = \frac{d}{dt} e^{\lambda(e^t-1)} = e^{\lambda(e^t-1)}\lambda e^t.$$

   We obtain $\mathbb{E}X$ by evaluating this at zero, which gives

$$M_X^{(1)}(0) = \lambda.$$

   ii. $\mathbb{E}X^2$.

   We have

$$M_X^{(2)}(t) = \left(\frac{d}{dt}\right)^2 e^{\lambda(e^t-1)} = e^{\lambda(e^t-1)}\lambda e^t \cdot \lambda e^t + e^{\lambda(e^t-1)}\lambda e^t.$$

   We obtain $\mathbb{E}X^2$ by evaluating this at zero, which gives

$$M_X^{(2)}(0) = \lambda^2 + \lambda.$$

iii. $\operatorname{Var} X$.

By the useful expression we have

$$\operatorname{Var} X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

(c) Let $Y = 3X + 1$. Give the mgf of Y and state whether Y has a Poisson distribution.

We have

$$M_Y(t) = M_{3X+1}(t) = e^t M_X(3t) = e^t e^{\lambda(e^{3t}-1)}.$$

This is not the mgf of a Poisson distribution, so Y does *not* have a Poisson distribution.

2. Let $X \sim \text{Uniform}(0, \theta)$ distribution.

(a) Show that the mgf of X can be written as

$$1 + \frac{t\theta}{2} + \frac{(t\theta)^2}{3!} + \frac{(t\theta)^3}{4!} + \frac{(t\theta)^4}{5!} + \dots$$

Hint: Make use of the series representation

$$e^a = \begin{cases} \sum_{i=0}^{\infty} a^i/i!, & a \neq 0 \\ 1, & a = 0. \end{cases}$$

For $t \neq 0$, the mgf of X can be written

$$\begin{aligned} M_X(t) &= \mathbb{E}e^{tX} \\ &= \int_0^\theta e^{tx} \cdot \frac{1}{\theta} dx \\ &= \frac{1}{\theta} \int_0^\theta \sum_{i=0}^{\infty} \frac{(tx)^i}{i!} dx \\ &= \frac{1}{\theta} \sum_{i=0}^{\infty} \frac{t^i}{i!} \int_0^\theta x^i dx \\ &= \frac{1}{\theta} \sum_{i=0}^{\infty} \frac{t^i}{i!} \frac{\theta^{i+1}}{(i+1)} \\ &= \sum_{i=0}^{\infty} \frac{(t\theta)^i}{(i+1)!} \\ &= 1 + \frac{t\theta}{2} + \frac{(t\theta)^2}{3!} + \frac{(t\theta)^3}{4!} + \frac{(t\theta)^4}{5!} + \dots \end{aligned}$$

(b) Identify the distribution of the rv $Y = X/\theta$ by finding its mgf.

We have
$$M_Y(t) = M_{X/\theta}(t) = M_X(t/\theta) = 1 + \frac{t}{2} + \frac{t^2}{3!} + \frac{t^3}{4!} + \frac{t^4}{5!} + \dots$$
which is the mgf of the Uniform$(0,1)$ distribution, so $Y \sim \text{Uniform}(0,1)$.

3. Let $X \sim \text{Gamma}(2,2)$. *Hint: Make use of the* `pgamma()` *and* `qgamma()` *functions in R.*

(a) Give $P(X > 2)$.

Using R, we have `1 - pgamma(2,2,scale = 2)` $= 0.7357589$.

(b) Give the median of X.

Using R, we have `qgamma(1/2,2,scale = 2)` $= 3.356694$.

(c) Find the mgf of the rv $Y = 2X - 4$ and state whether Y has a Gamma distribution.

We have
$$M_Y(t) = M_{2X-4}(t) = e^{-4t}M_X(2t) = e^{-4t}(1-4t)^{-2}.$$
This is not the mgf of a Gamma distribution, so Y does *not* have a Gamma distribution.

(d) Find $P(Y < 1)$.

We have

$P(Y < 1) = P(2X - 4 < 1) = P(X < (1+4)/2) =$ `pgamma(5,2,scale=2)` $= 0.7127025$.

(e) Find the mgf of the rv $W = 2X$ and state whether W has a Gamma distribution.

We have
$$M_W(t) = M_{2X}(t) = M_X(2t) = (1-4t)^{-2},$$
which is the mgf of the Gamma$(2,4)$ distribution, so $W \sim \text{Gamma}(2,4)$.

(f) Find $P(1 < W < 2)$.

We have

$P(1 < W < 2) = P(W \leq 2) - P(W \leq 1)$
$=$ `pgamma(2,2,scale=4) - pgamma(1,2,scale=4)`
$= 0.06370499$.

4. Find the quantile function $Q_X(\theta) : (0, 1) \to \mathcal{X}$ for each of the following random variables (*Hint*: Set up the equation $F_X(q) = \theta$ and solve for q):

(a) $X \sim \text{Exponential}(\lambda)$.

Note that X has cdf

$$F_X(x) = 1 - e^{-x/\lambda} \quad \text{for} \quad x \leq 0.$$

So setting $F_X(q) = \theta$ and solving for q, we obtain

$$Q_X(\theta) = -\lambda \log(1 - \theta) \quad \text{for} \quad \theta \in (0, 1).$$

(b) X having cdf given by

$$F_X(x) = \frac{1}{[1 + e^{-\tau(x-\mu)}]^{1/\nu}}, \quad -\infty < x < \infty.$$

for some $\tau > 0$, $\nu > 0$, and $\mu \in \mathbb{R}$.

Setting $F_X(q) = \theta$ and solving for q, we obtain

$$Q_X(\theta) = \mu + \frac{1}{\tau} \log\left(\frac{1 - \theta^\nu}{\theta^\nu}\right) \quad \text{for} \quad \theta \in (0, 1).$$

5. Consider the set of data points

$$0.27 \quad -0.63 \quad 0.87 \quad 1.73 \quad 0.02 \quad 0.37 \quad -1.31 \quad 0.74 \quad 0.04 \quad -1.05.$$

(a) Find the θ-quantile of the empirical distribution of these data points for $\theta = (i - 0.5)/10$, for $i = 1, \dots, 10$.

These are the quantiles $0.05, 0.15, \dots, 0.95$, and they are given by the sorted data values, since $x_{(\lceil 10*0.05 \rceil)}, \dots, x_{(\lceil 10*0.95 \rceil)} = x_{(1)}, \dots, x_{(10)}$. The sorted data values are

$$-1.31 \quad -1.05 \quad -0.63 \quad 0.02 \quad 0.04 \quad 0.27 \quad 0.37 \quad 0.74 \quad 0.87 \quad 1.73$$

(b) Give the θ-quantile for $\theta = (i - 0.5)/10$, for $i = 1, \dots, 10$ of the $\text{Normal}(0, 1)$ distribution.

We obtain these as $\Phi^{-1}((i - 0.5)/10)$, for $i = 1, \ldots, 10$, where Phi^{-1} is the inverse of the standard Normal cdf. The R code `round(qnorm( (c(1:10) - .5)/ 10),3)` gives the values

$$-1.645 \quad -1.036 \quad -0.674 \quad -0.385 \quad -0.126 \quad 0.126 \quad 0.385 \quad 0.674 \quad 1.036 \quad 1.645$$

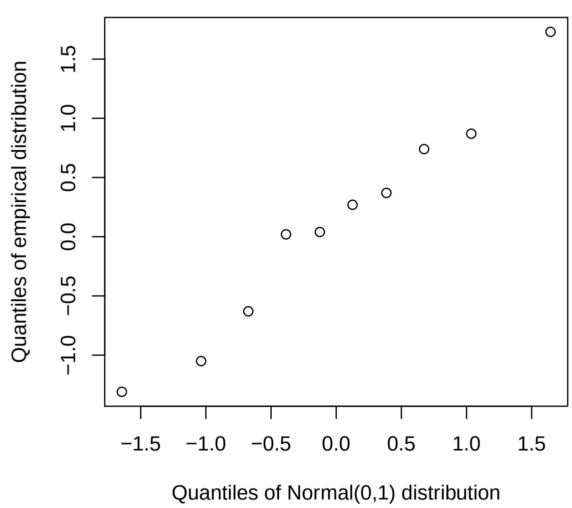
(c) Make a plot of the empirical distribution quantiles (on the vertical axis) versus the Normal$(0, 1)$ quantiles. *Use whatever software you want. Print the plot or take a picture of it on your screen.*

The R code

```
emp.quantiles <- c(-1.31, -1.05, -0.63, 0.02, 0.04, 0.27, 0.37, 0.74, 0.87, 1.73)
theo.quantiles <- qnorm( (c(1:10) - .5)/ 10)

plot(emp.quantiles ~ theo.quantiles,
     ylab = "Quantiles of empirical distribution",
     xlab = "Quantiles of Normal(0,1) distribution")
```

produces the plot



(d) You should see that the points fall roughly along a straight line. What is your interpretation of this?

That the points fall roughly on a straight line indicates that the quantiles of the empirical distribution are close to the corresponding quantiles of the Normal$(0, 1)$ distribution.

Therefore, we might assume that the data values were sampled from the Normal$(0, 1)$ distribution.

6. Let $X \sim \text{Binomial}(3, 1/2)$.
   (a) Make a drawing of the cdf of X.
   (b) Find the θ-quantile of X for all the values $\theta = 2/16, 3/16, 8/16, 9/16, 15/16$.

We have

θ	2/16	3/16	8/16	9/16	15/16
q_θ	0	1	1	2	3

Optional (do not turn in) problems for additional study from Wackerly, Mendenhall, Scheaffer, 7th Ed.*:

- 3.145, 3.146, 3.147, 3.148, 3.149, 3.150, 3.153, 3.154
- 4.42, 4.61
- 4.136, 4.139, 4.144, 4.145

*Ignore all references in the textbook to applets and just use R to compute probabilities that cannot be computed by hand.