MTH5129 Probability & Statistics II

Coursework 5

1. Let X have the probability density function

$$f_X(x) = \begin{cases} 4x^3, & \text{for } 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

Find the probability density function of $Y = 2X - 1$ using the transformation of random variables method.

Solution: This can be proved in many ways, e.g. using the cumulative distribution function method or the transformation of random variables method. We actually used the former way in last week's coursework). We use the latter way here.

Using the direct transformation of random variables method, the inverse of $y = 2x - 1$ is $g^{-1}(y) = (y + 1)/2$ and the domain is $-1 < y < 1$ (the initial range corresponding to $0 < x < 1$). We have

$$\frac{d}{dy} g^{-1}(y) = \frac{1}{2}.$$

The theorem of Transformation of random variables implies that

$$f_Y(y) = \begin{cases} 4\left(\frac{y+1}{2}\right)^3 \frac{1}{2} = 2\left(\frac{y+1}{2}\right)^3 = \frac{1}{4}(y+1)^3, & \text{for } -1 < y < 1 \\ 0, & \text{otherwise.} \end{cases}$$

2. Suppose that X_1 and X_2 have joint probability density function

$$f(x_1, x_2) = \begin{cases} 8x_1x_2, & 0 < x_1 < x_2 < 1 \\ 0, & \text{otherwise.} \end{cases}$$

What is the probability density function of $Y_1 = X_1/X_2$?
Hint: Define an additional variable, e.g. $Y_2 = X_2$.

Solution: In this case, we need another variable, thus we choose $Y_2 = X_2$, as this choice will allow us to find the inverse easily. Any other choice of Y_2 is also acceptable.

The inverse transformation is $x_1 = y_1 y_2$ and $x_2 = y_2$.

The Jacobian is

$$J = \begin{vmatrix} y_2 & y_1 \\ 0 & 1 \end{vmatrix} = y_2.$$

The support of (X_1, X_2) is given by $A = \{0 < x_1 < x_2 < 1\}$ which implies that the support of (Y_1, Y_2) is therefore $B = \{0 < y_1 y_2 < y_2 < 1\}$. By rearranging this, we get that $B = \{0 < y_1 < 1,\ 0 < y_2 < 1\}$.

So, using the direct transformation of two random variables method

$$f_{Y_1,Y_2}(y_1, y_2) = \begin{cases} [8(y_1\, y_2)y_2]\, y_2 = 8y_1\, y_2^3, & \text{for } (y_1, y_2) \in B \\ 0, & \text{otherwise.} \end{cases}$$

We now find the marginal probability density function $f_{Y_1}(y_1)$ of Y_1.

We firstly have $f_{Y_1}(y_1) = 0$ for $y_1 \leq 0$ and $y_1 \geq 1$, due to the support B of the joint probability density function $f_{Y_1,Y_2}(y_1, y_2)$ obtained above.

Moreover, for $0 < y_1 < 1$, we have

$$f_{Y_1}(y_1) = \int_0^1 8y_1 y_2^3\, dy_2 = 8y_1 \left[\frac{y_2^4}{4}\right]_0^1 = 2y_1$$

therefore we have

$$f_{Y_1}(y_1) = \begin{cases} 2y_1, & \text{for } 0 < y_1 < 1 \\ 0, & \text{otherwise.} \end{cases}$$

3. Suppose X has a normal distribution, $N(\mu, \sigma^2)$, find the moment generating function of X and then deduce its mean and variance.

Solution:

$$\begin{aligned} M_X(t) &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}\, dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}[x^2 - 2\mu x + \mu^2 - 2\sigma^2 tx]\right\}\, dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}[x^2 - 2x(\mu + \sigma^2 t) + \mu^2]\right\}\, dx \end{aligned}$$

Now we complete the square in x:

$$\begin{aligned} [x^2 - 2x(\mu + \sigma^2 t) + \mu^2] &= [x - (\mu + \sigma^2 t)]^2 + \mu^2 - (\mu + \sigma^2 t)^2 \\ &= [x - (\mu + \sigma^2 t)]^2 - (2\mu\sigma^2 t + \sigma^4 t^2) \end{aligned}$$

and so as the final bracket does not depend on x we can take it outside the integral to give

$$M_X(t) = \exp\left(\frac{2\mu\sigma^2 t + \sigma^4 t^2}{2\sigma^2}\right) \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}[x - (\mu + \sigma^2 t)]^2\right\}\, dx$$

Now the function inside the integral is the probability density function of a $N(\mu + \sigma^2 t, \sigma^2)$ random variable and so is equal to one. Thus

$$M_X(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right).$$

Now differentiating the moment generating function we find

$$M'(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)(\mu + \sigma^2 t),$$

thus $M'(0) = \mu = E[X]$. Moreover,

$$M''(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)(\mu + \sigma^2 t)^2 + \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)\sigma^2,$$

thus $M''(0) = \mu^2 + \sigma^2$ and hence $\mathrm{Var[X]} = \mu^2 + \sigma^2 - (\mu)^2 = \sigma^2$.

4. Let X be a $Bin(n, p)$ random variable.

   (a) Find the moment generating function of X.

   (b) Find the expectation of X.

   (c) Find the variance of X.

   Solution:

   (a) The probabilities of a $Bin(n, p)$ random variable X are given by

   $$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

   for $x = 0, 1, \ldots, n$. The moment generating function is given by

   $$\begin{aligned} M_X(t) &:= E[e^{tX}] \\ &= \sum_{x=0}^{n} e^{tx}\, P(X = x) \\ &= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x (1-p)^{n-x} \\ &= (pe^t + 1 - p)^n \quad \text{(by the Binomial Theorem)} \end{aligned}$$

   Therefore

   $$M_X(t) = (pe^t + 1 - p)^n$$

(b) The expectation of X is given by

$$E[X] = M'_X(0) = n(pe^t + 1 - p)^{n-1} pe^t \Big|_{t=0} = np$$

(c) We also have

$$\begin{aligned} E[X^2] = M''_X(0) &= n(n-1)(pe^t + 1 - p)^{n-2} p^2 e^{2t} + n(pe^t + 1 - p)^{n-1} pe^t \Big|_{t=0} \\ &= n(n-1)p^2 + np \end{aligned}$$

hence the variance is given by

$$\mathrm{Var}(X) = n(n-1)p^2 + np - (np)^2 = np(1-p).$$

5. Suppose $Y_1, \ldots, Y_n$ are independent, normally distributed with mean $E[Y_i] = \mu_i$ and variance $\mathrm{Var[Y_i]} = \sigma_\mathrm{i}^2$. Prove that the sum

$$U = \sum_{i=1}^{n} \left(\frac{Y_i - \mu_i}{\sigma_i} \right)^2 \qquad \text{has a } \chi^2(n) \text{ distribution.}$$

Solution: Let

$$Z_i = \frac{Y_i - \mu_i}{\sigma_i},$$

for all $i = 1, \ldots, n$. We first have to prove that each Z_i^2 has a $\chi^2(1)$ distribution (Prove this! – answer is in your lecture notes). Therefore,

$$M_{Z_i^2}(t) = (1-2t)^{-1/2}.$$

The random variable we aim for can then be rewritten as

$$U = \sum_{i=1}^{n} Z_i^2.$$

Finally, since the Z_i's are independent, due to the X_i's being independent, we can calculate

$$\begin{aligned} M_U(t) &= \prod_{i=1}^{n} M_{Z_i^2}(t) \\ &= \frac{1}{(1-2t)^{n/2}} \\ &= \left(\frac{\frac{1}{2}}{\frac{1}{2} - t} \right)^{\frac{n}{2}} \end{aligned}$$

which is the moment generating function of a $Ga(n/2, 1/2)$ random variable, that is a $\chi^2(n)$ random variable.

Note: The above proves that the sum of n independent $\chi^2(1)$ random variables has a $\chi^2(n)$ distribution.

6. Suppose $X_1, X_2, \ldots, X_n$ are independent random variables with cumulant generating functions $K_{X_i}(t),\ i = 1, \ldots, n$. Let $Y = \sum_{i=1}^n X_i$ and find the cumulant generating function $K_Y(t)$ in terms of the $K_{X_i}(t),\ i = 1, \ldots, n$.

   Solution: We have

$$
\begin{aligned}
K_Y(t) := \ln M_Y(t) &= \ln\left(E[e^{tY}]\right) \\
&= \ln\left(E[e^{t\sum_{i=1}^n X_i}]\right) \\
&= \ln\left(E[e^{tX_1}e^{tX_2}\cdots e^{tX_n}]\right) \\
&= \ln\left(E[e^{tX_1}]E[e^{tX_2}]\cdots E[e^{tX_n}]\right) \qquad \text{(by independence)} \\
&= \ln\left(M_{X_1}(t)M_{X_2}(t)\cdots M_{X_n}(t)\right) \\
&= \sum_{i=1}^n \ln M_{X_i}(t) = \sum_{i=1}^n K_{X_i}(t)\,.
\end{aligned}
$$

7. Suppose that X is a non-negative random variable with mean μ. Prove that the median is at most 2μ. (The median is a value m with $P(X \geq m) \geq 1/2$ and $P(X \leq m) \geq 1/2$: i.e. it is the "middle value".)

   Solution: Suppose m is the median: then we know that $P(X \geq m) \geq 1/2$. But by Markov's Inequality we know that

$$\frac{1}{2} \leq P(X \geq m) \leq \frac{E(X)}{m} = \frac{\mu}{m}.$$

   Rearranging we get that $m \leq 2\mu$ as claimed.

8. Suppose that I toss a fair coin 100 times. Prove that the probability I get more than or equal to 60 heads or less than or equal to 40 heads is at most 1/4.

   Solution: Let X be the number of heads. Then $\mu = E(X) = 50$ and $\sigma^2 = \mathrm{Var}(X) = 25$. Hence, by Chebyshev's inequality

$$
\begin{aligned}
P(X \leq 40 \text{ or } X \geq 60) &= P(|X - 50| \geq 10) \\
&= P(|X - E(X)| \geq 10) \leq \frac{\sigma^2}{100} = 1/4.
\end{aligned}
$$