Oct. 7-11

Stat 330 - Tutorial 5

This week's tutorial is about transformations. This topic is one of the more technically challenging parts of the course so, much like with combinatorics, it will help to do lots of practice. I will discuss the foundations ~~and~~ and show you some ~~techniques~~ strategies, but it's difficult to grasp the technique unless you do lots of practice.

Discrete Distributions

Transformations are more straightforward when we're working with discrete distributions, so I'm going to spend less time here. Let X be a discrete random variable and $Y=g(X)$, where g is some invertible function. We can get the distribution of Y directly from the distribution of X.

$$P(Y=y)=p(g(X)=y)=p(X=g^{-1}(y))$$

~~As~~ As long as we have a formula for $p(X=x)$, we can get a formula for $P(Y=y)$ by plugging in $x=g^{-1}(y)$.

eg. 1: Let X be the outcome of a single fair dice roll ~~and~~ and $g(x)=2x-1$. Let $Y=g(X)$. What is the distribution of Y?

In practice, you would probably do this question with a table of values. Let's use the transformations strategy anyway. The ~~pdf~~ PMF of X is $p(X=x)=\frac{1}{6}I(x=1,\dots,6)$. The inverse of our transformation is $g^{-1}(y)=\frac{y+1}{2}$. Therefore, the PMF of Y is

$$P(Y=y)=p(X=g^{-1}(y))=p(X=\tfrac{y+1}{2})=\frac{1}{6}I(\tfrac{y+1}{2}=1,\dots,6)=\frac{1}{6}I(y=1,3,\dots,11)$$

◇

Continuous Distributions

Now let X be a continuous R.V. ~~and let g be~~ and let $Y=g(X)$. We again assume that g is invertible, but ~~also require that~~ will now also require that g be increasing. We want to get the distribution of Y, so let's start with the CDF.

$$F_Y(y) = P(Y \le y) = P(g(X) \le y) = P(X \le g^{-1}(y)) = F_X(g^{-1}(y))$$

If we have a nice formula for the CDF of X (like when $X \sim \exp$), we're in good shape. We can evaluate $F_X(x)$ at $x = g^{-1}(y)$ and hope that we recognize the outcome as a function of y. ~~Most distributions are~~ Most continuous distributions are described by their density however, so getting the density of Y would make our life easier. Calculating the density of Y takes a bit more care. From the definition, we get

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X(g^{-1}(y))$$

This is where we need to be careful. We know that the density of X is $f_X(x) = \frac{d}{dx} F_X(x)$. Let's use this fact with the chain rule to expand our expression on the right. Let $z = g^{-1}(y)$

$$\frac{d}{dy} F_X(g^{-1}(y)) = \frac{d}{dz} F_X(z) \cdot \frac{d}{dy} g^{-1}(y) = f_X(z) \cdot \frac{d}{dy} g^{-1}(y) = f_X(g^{-1}(y)) \cdot \frac{d}{dy} g^{-1}(y)$$

We now have the transformation formula for continuous random variables with invertible increasing transformations:

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \frac{d}{dy} g^{-1}(y)$$

This formula isn't the most intuitive, so let's do an example.

e.g.2: Let $X \sim \exp(\lambda)$ and $Y = 2X$. Find the distribution of Y.

First, let's write $Y = g(X)$, where $g(x) = 2x$. Then g is increasing and invertible, so we can proceed. Next, note that $f_X(x) = \lambda e^{-\lambda x}$ and $g^{-1}(y) = y/2$, so

$$f_X(g^{-1}(y)) = \lambda e^{-\lambda x/2} \quad \text{and} \quad \frac{d}{dy} g^{-1}(y) = \frac{1}{2}$$

Therefore, the density of Y is

$$f_Y(y) = \left(\lambda e^{-\lambda x/2}\right) \cdot \left(\frac{1}{2}\right) = \left(\frac{\lambda}{2}\right) e^{-(\lambda/2) y}$$

This is the density of an $\exp(\lambda/2)$ random variable, so $Y \sim \exp(\lambda/2)$.

$\diamond$

The formula $f_X(g^{-1}(y)) \cdot \frac{d}{dy} g^{-1}(y)$ is messy and hard to remember. I find it easier to think of $g(x)$ as $Y(x)$, since g maps from X-values to Y-values. Similarly, I think of $g^{-1}(y)$ as $X(y)$, since g^{-1} maps from Y back to X. Written this way, our formula for $f_Y(y)$ is

$$f_X(X(y)) \cdot \frac{d}{dy} X(y)$$

The first term makes more sense now (to me at least). It is telling us to find the X-value corresponding to our y, $X(y)$, and plug this X-value into the density for X.
The second term is always going to be a bit messy, since it is a technical correction for the fact that we're working with densities instead of with actual probabilities or events.

There's much more we could do here, but I want to leave time to answer questions about the midterm. Let's end on a harder example.

e.g. 3: Let $X \sim N(0,1)$. Let $Y = X^2$. Show that $Y \sim \chi^2(1)$. That is, show that the density of Y is $f_Y(y) = \frac{1}{\Gamma(\frac{1}{2})2^{1/2}} y^{-1/2} e^{-y/2} = \frac{1}{\sqrt{2\pi y}} e^{-y/2}$

Let's start by writing $Y = g(X)$, where $g(x) = x^2$. Note that g is neither invertible nor increasing, so we can't apply our methodology directly. However, if X were restricted to positive values then g would be good and we would have $g^{-1}(y) = \sqrt{y}$. This is probably going to be important, so let's define $h(y) = \sqrt{y}$. Our method doesn't apply here, but it might be helpful to see what we would have gotten if it did. Let's write

$$k(y) = f_X(h(y)) \cdot \frac{d}{dy} h(y) =: a(y) \cdot b(y)$$

Then $a(y) = \frac{1}{\sqrt{2\pi}} e^{-h(y)^2/2} = \frac{1}{\sqrt{2\pi}} e^{-y/2}$ $\qquad b(y) = \frac{d}{dy} h(y) = \frac{d}{dy}\sqrt{y} = \frac{1}{2\sqrt{y}}$

Putting these together, we get $k(y) = \frac{1}{\sqrt{2\pi}} e^{-y/2} \cdot \frac{1}{2\sqrt{y}} = \frac{1}{2\sqrt{2\pi y}} e^{-y/2}$

This is actually really close to what we want. Now let's do something that actually applies in this question: the CDF method.

$$F_Y(y) = P(Y \le y) = P(X^2 \le y) = P(-\sqrt{y} \le X \le \sqrt{y}) = P(X \le \sqrt{y}) - P(X < -\sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

From the symmetry of the normal distribution, we get that $F_X(-\sqrt{y}) = P(X > \sqrt{y}) = 1 - F_X(\sqrt{y})$.

Therefore, $F_Y(y) = F_X(\sqrt{y}) - [1 - F_X(\sqrt{y})] = 2F_X(\sqrt{y}) - 1 = 2F_X(h(y)) - 1$

Finally, we get the density of Y by differentiating its CDF.

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy}[2F_X(h(y)) - 1] = 2\left[f_X(h(y)) \cdot \frac{d}{dy} h(y)\right] = 2k(y) = \frac{1}{\sqrt{2\pi y}} e^{-y/2}$$

This is the density of a $\chi^2(1)$ random variable, so we are done. $\diamond$