STA 623 – Fall 2013 – Dr. Charnigo

Section 2.1: Distributions of functions of a random variable

Probabilities for functions of a random variable. Let X be a random variable. Put $Y := g(X)$ for some function g such that, for any set $B \in \mathcal{B}^1$, we have $g^{-1}(B) := \{x \in \mathbb{R} : g(x) \in B\} \in \mathcal{B}^1$. (This condition is satisfied for all g of practical interest.) Then, for any $B \in \mathcal{B}^1$, we define $P(Y \in B) := P(X \in g^{-1}(B))$.

Example (probabilities for functions of a random variable). Suppose that X has probability density function $\exp[-2|x|]$. Put $Y := g(X) := X^2$ and $B := [0, y]$ for a generic nonnegative real y. Then $g^{-1}(B) =$
and $P(Y \leq y) =$

Discrete and continuous cases. If X is a discrete random variable, then so is Y. In this case, taking $B := \{y\}$ for a generic $y \in \mathbb{R}$ shows that

$$P(Y = y) = \sum_{x \in g^{-1}(\{y\}) : P(X=x)>0} P(X = x).$$

If X is a continuous random variable, the same may or may not be true of Y. For instance, let X have probability density function $1_{\{x>0\}} \exp[-x]$. Then $Y := X^2$ is a continuous random variable, $Y := \lfloor X \rfloor$ is a discrete random variable (here $\lfloor X \rfloor$, read "the floor of X", is the largest integer less than or equal to X), and $Y := \min\{X, 2\}$ is neither a discrete random variable nor a continuous random variable.

Example (discrete and continuous cases). Let X have probability mass function $(1/2)^x$ for $x \in \{1, 2, \ldots\}$. Let $Y := \lfloor X \rfloor \bmod 2$, so that $Y = 1$ when $\lfloor X \rfloor$ is odd and $Y = 0$ otherwise. We have $P(Y = 0) =$

Monotonicity of the transforming function. Let $f_X(x)$ denote the probability mass (if X is a discrete random variable) or density (if X is a continuous random variable) function of X. Let $\mathcal{X} := \{x \in \mathbb{R} : f_X(x) > 0\}$, which we refer to as the support of X, and let $\mathcal{Y} := \{y \in \mathbb{R} : y = g(x) \text{ for some } x \in \mathcal{X}\}$.

First suppose that $g(x)$ is strictly increasing, in that $u > v$ implies $g(u) > g(v)$. Then, for any $y \in \mathcal{Y}$ there exists a unique $x \in \mathcal{X}$ such that $g(x) = y$. We refer to this x as $g^{-1}(y)$. Since $g(X) \leq y$ is equivalent to $X \leq g^{-1}(y)$, we have

$$F_Y(y) = F_X(g^{-1}(y)),$$

where F_Y and F_X denote the cumulative distribution functions of Y and X respectively. In addition, if X is a continuous random variable, $f_X(x)$ is continuous on $\mathcal{X}$, and $g^{-1}(y)$ has a continuous derivative on $\mathcal{Y}$, then Y is a continuous random variable with probability density function

$$f_Y(y) = f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y) 1_{\{y \in \mathcal{Y}\}}.$$

Next suppose that $g(x)$ is strictly decreasing, in that $u > v$ implies $g(u) < g(v)$. Again, for any $y \in \mathcal{Y}$ there exists a unique $x \in \mathcal{X}$ such that $g(x) = y$. We refer to this x as $g^{-1}(y)$. Since $g(X) \leq y$ is equivalent to $X \geq g^{-1}(y)$, we have

$$F_Y(y) = 1 - F_X(g^{-1}(y)) + P(X = g^{-1}(y)).$$

(Where does the last term above come from?) In addition, if X is a continuous random variable, $f_X(x)$ is continuous on $\mathcal{X}$, and $g^{-1}(y)$ has a continuous derivative on $\mathcal{Y}$, then Y is a continuous random variable with probability density function

$$f_Y(y) = -f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y)1_{\{y\in\mathcal{Y}\}}.$$

There is also a formula in your textbook for $f_Y(y)$ when, among other conditions, $\mathcal{X}$ can be partitioned into finitely many intervals on each of which $g(x)$ is monotone (i.e., either strictly increasing or strictly decreasing). Personally I think that deriving $F_Y(y)$ from first principles and then differentiating in y is easier than trying to remember the textbook formula.

Example (monotonicity of the transforming function). Suppose that X has probability density function $\exp[-x]1_{\{x>0\}}$ with corresponding cumulative distribution function $(1-\exp[-x])1_{\{x>0\}}$. Put $g(x) := 1-\exp[-x]$, which is strictly increasing and which maps $\mathcal{X} = (0,\infty)$ to $\mathcal{Y} = (0,1)$. We have $g^{-1}(y) = -\log(1-y)$, from which we find that a probability density function for Y is In fact, this illustrates a general result called the probability integral transformation: whenever $g(x) = F_X(x)$ for a continuous random variable X, we obtain this probability density function for Y. (The general result is more challenging to verify because there exist continuous random variables whose cumulative distribution functions are not strictly increasing.)

Suppose that X has probability density function $\exp[-2|x|]$. Put $g(x) := x^2$, which is clearly not monotone. However, since we found $P(Y \leq y)$ earlier, we can differentiate in y to obtain

an expression that is valid for all $y \in (0,\infty)$. We can define $f_Y(y)$ to be zero for all $y \in (-\infty, 0]$, and then $f_Y(y)$ will be a probability density function for Y.