MATH 2030 3.00MW – Elementary Probability Course Notes Part III: Random Variables

Tom Salisbury – salt@yorku.ca

York University, Dept. of Mathematics and Statistics
Original version April 2010. Thanks are due to
E. Brettler, V. Michkine, and R. Shieh for many corrections

January 18, 2018

Discrete Distributions

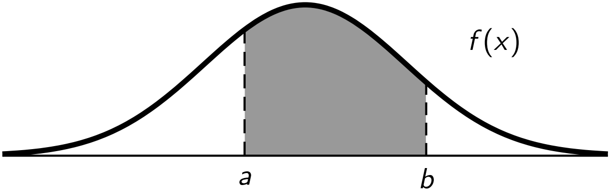
- Recall that a r.v. has a *discrete distribution* if there are only finitely or countably many values it can take.
- In that case, the distribution is given by a table of values x_i and probabilities $p_i = P(X = x_i)$.

x	x_1	x_2	x_3	$\cdots$
$P(X = x)$	p_1	p_2	p_3	$\cdots$

- Conditions the x_i and p_i must satisfy:
  the x_i are distinct; the $p_i \geq 0$ and $\sum p_i = 1$.
  [ie. any x_i and p_i satisfying these conditions define a discrete distribution]

Continuous Distributions

- A r.v. X has a *continuous distribution* (more properly, an *absolutely continuous distribution*) if there is a function $f(t)$ such that $P(X \in I) = \int_I f(t)\,dt$ for every interval $I = [a, b]$.



- This f is called the *probability density function* of X.
- The conditions the density must satisfy are that: $f \geq 0$ and $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

Continuous Distributions

- **Eg:** We'll study the *Normal Distribution* later, with $f(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$
- **Eg:** X has a *uniform distribution* on $[a, b]$ if $f(x) = c$ for $x \in [a, b]$ and $f(x) = 0$ otherwise, where c is some constant and $a < b$. Our conditions then $\Rightarrow c = \frac{1}{b-a}$.
- For example, if X is uniform on $[1, 4]$ then

$$f(x) = \begin{cases} 0, & x < 1 \\ \frac{1}{3}, & 1 < x < 4 \\ 0, & x > 4. \end{cases}$$

  Therefore $P(X > 2) = \int_2^\infty f(x)\,dx = \int_2^4 \frac{1}{3}\,dx + \int_4^\infty 0\,dx = \frac{2}{3}$.
- X has a continuous distribution $\Rightarrow P(X = x) = 0\ \forall x$. So a density can be changed at finitely many points without altering any probabilities. *Densities aren't unique.* (& it doesn't matter how/if I fill in $x = 1, 4$ values for the above f.)

Distribution functions

The *cumulative distribution function* (cdf) of X is

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= \begin{cases} \sum_{i: x_i \leq x} p_i, & \text{dist'n of } X \text{ is discrete} \\ \int_{-\infty}^{x} f(t)\, dt, & \text{dist'n of } X \text{ is continuous} \end{cases} \end{aligned}$$

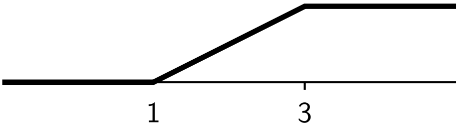
Eg: Let $X \sim$ uniform on $[1, 3]$. Find the cdf.

- We have $f(x) = \begin{cases} 0, & x < 1 \\ \frac{1}{2}, & 1 < x < 3 \\ 0, & x > 3. \end{cases}$
- We must compute $F(x) = \int_{-\infty}^{x} f(t)\, dt$
- If $x < 1$, $F(x) = \int_{-\infty}^{x} 0\, dt = 0$

Distribution functions

- If $1 \leq x < 3$, $F(x) = \int_{-\infty}^{1} 0\,dt + \int_{1}^{x} \frac{1}{2}\,dt = 0 + \frac{x-1}{2} = \frac{x-1}{2}$
- If $3 \leq x$, $F(x) = \int_{-\infty}^{1} 0\,dt + \int_{1}^{3} \frac{1}{2}\,dt + \int_{3}^{x} 0\,dt = 0 + 1 + 0 = 1$

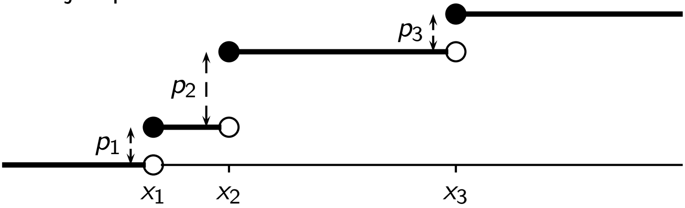
$$\text{Therefore} \quad F(x) = \begin{cases} 0, & x < 1 \\ \frac{x-1}{2}, & 1 \leq x < 3 \\ 1, & 3 \leq x. \end{cases}$$



Recovering information from cdf's

We've just seen how to compute cdf's. But the cdf is useful because it encodes probabilities or densities for r.v.'s, so we need to be able to recover that information from the cdf.

- If X has a continuous distribution, we recover the density from $f = F'$ (typically at all but finitely many points, where F might not be differentiable).
- If X has a discrete distribution, then F is a "staircase". The values x_i are where F jumps. The probabilities p_i are the sizes of the jumps.



Characterization of cdf's

If F is the cdf of a r.v. X then

- $F \uparrow$, ie. $x \leq y \Rightarrow F(x) \leq F(y)$
- $\lim_{x\to\infty} F(x) = 1$
- $\lim_{x\to-\infty} F(x) = 0$
- F is right continuous,
  ie. $\forall x \; F(x) = F(x+) = \lim_{y\to x, y<x} F(y)$.

In fact, these conditions characterize cdf's:

Theorem F is a cdf $\Leftrightarrow$ the above 4 conditions hold

We won't prove this. But what $\Leftarrow$ means is that if $F : \mathbb{R} \to \mathbb{R}$ is any function satisfying the 4 conditions above, then there exists a model $(\Omega, \mathcal{F}, P)$ and a r.v. X such that F is the cdf of X.
Eg: A cdf that is neither discrete nor continuous (of a *mixture* of discrete and continuous dist'ns)

Probabilities from cdf's

- $P(X \leq b) = F(b)$
- $P(X \in (a, b]) = F(b) - F(a)$
- $P(X > a) = 1 - F(a)$
- $P(X < b) = F(b-)$
- $P(X \geq a) = 1 - F(a-)$
- $P(X = a) = F(a) - F(a-)$
- $P(X \in (a, b)) = F(b-) - F(a)$
- $P(X \in [a, b]) = F(b) - F(a-)$
- $P(X \in [a, b)) = F(b-) - F(a-)$

And of course, if X has a continuous distribution then $F(x-) = F(x)\ \forall x$, and the above simplify.
We'll use these a lot for standard normal distributions, where F is given by a table such as Appendix 5.

Eg: Uniform random points

Problem: Pick a point (X, Y) uniformly at random from $C = [0,1] \times [0,1]$. Let $Z = X + Y$. Find the density of Z.

We'll solve this by finding the cdf $F(z)$ of Z, and then differentiating.
First we need to interpret "uniformly" in this context:
Definition: Picking a point (X, Y) *uniformly at random* from a set C means that $P((X, Y) \in A) = \dfrac{\text{area}(A \cap C)}{\text{area}(C)}$, for $A \subset \mathbb{R}^2$.

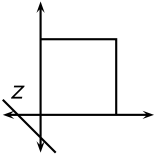
So let $A_z = \{(x, y) \mid 0 \leq x, y \leq 1, x + y \leq z\}$. Then

$$\begin{aligned} F(z) &= P(X + Y \leq z) = P((X, Y) \in A_z) \\ &= \frac{\text{area}(A_z)}{\text{area}(C)} = \frac{\text{area}(A_z)}{1} = \text{area}(A_z). \end{aligned}$$

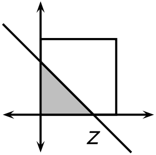
We have the following cases:

Uniform random points

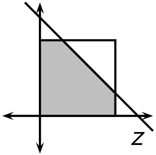
$$\text{area}(A_z) = 0$$



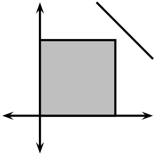
$$\text{area}(A_z) = z^2/2$$



$$\text{area}(A_z) = 1 - (2-z)^2/2$$



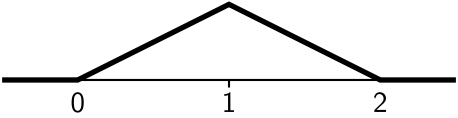
$$\text{area}(A_z) = 1$$



Uniform random points

Therefore $F(z) = \begin{cases} 0, & z < 0 \\ \frac{z^2}{2}, & 0 \le z < 1 \\ 1 - \frac{(2-z)^2}{2}, & 1 \le z < 2 \\ 1, & 2 \le z \end{cases}$

and so $f(z) = F'(z) = \begin{cases} 0, & z < 0 \\ z, & 0 < z < 1 \\ 2 - z, & 1 < z < 2 \\ 0, & 2 < z \end{cases}$



Transformations

- If we know the density $f(x)$ of X, and $Y = g(X)$, how do we find the density $h(y)$ of Y?
- This is easiest if g is one-to-one (ie strictly $\uparrow$ or $\downarrow$)

Problem: Find the density $h(y)$ of Y, if X is uniform on $[1, 3]$ and $Y = X^3$. (A slightly different version than the one I did in class)

As in the last eg, we'll find the cdf $H(y)$ of Y and differentiate. Given y, let $x = y^{1/3}$ so $x^3 = y$. Since $g(x) = x^3$ is strictly $\uparrow$, $H(y) = P(Y \le y) = P(X^3 \le x^3) = P(X \le x) = F(x) = F(y^{1/3})$. We worked out F before, and substituting gives

$$H(y) = \begin{cases} 0, & y < 1 \\ \frac{y^{1/3}-1}{2}, & 1 \le y < 27 \\ 1, & 27 \le y \end{cases} \text{ so } h(y) = \begin{cases} 0, & y < 1 \\ \frac{1}{6y^{2/3}}, & 1 < y < 27 \\ 0, & 27 < y \end{cases}$$

Transformations

Problem: Find the density $h(y)$ of Y, if X is uniform on $[-1, 2]$ and $Y = X^2$.

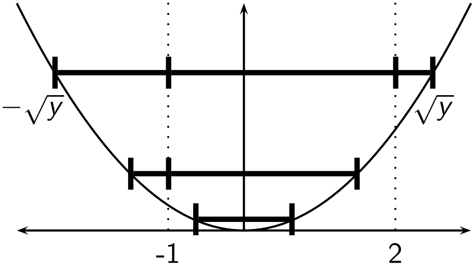
- Now $g(x) = x^2$ is not 1-1, which complicates the analysis.
- As before, we first find the cdf.

$$H(y) = P(Y \leq y) = P(X^2 \leq y)$$
$$= \begin{cases} 0, & y < 0 \\ P(|X| \leq \sqrt{y}), & y \geq 0 \end{cases}$$

- And for $y \geq 0$,
  $P(|X| \leq \sqrt{y}) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} f(x)\, dx$.
  But to work the latter out, we need to break into cases.

Transformations

- The density of X is $f(x) = \begin{cases} \frac{1}{3}, & -1 < x < 2 \\ 0, & \text{otherwise.} \end{cases}$
- If $0 \le y < 1$, $\int_{-\sqrt{y}}^{\sqrt{y}} f(x)\,dx = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{3}\,dx = 2\sqrt{y}/3$.
- If $1 \le y < 2^2 = 4$,
  $\int_{-\sqrt{y}}^{\sqrt{y}} f(x)\,dx = \int_{-\sqrt{y}}^{-1} 0\,dx + \int_{-1}^{\sqrt{y}} \frac{1}{3}\,dx = (\sqrt{y}+1)/3$.
- If $4 \le y$,
  $\int_{-\sqrt{y}}^{\sqrt{y}} f(x)\,dx = \int_{-\sqrt{y}}^{-1} 0\,dx + \int_{-1}^{2} \frac{1}{3}\,dx + \int_{2}^{\sqrt{y}} 0\,dx = 1$.



Transformations

- Therefore

$$H(y) = \begin{cases} 0, & y < 0 \\ \frac{2\sqrt{y}}{3}, & 0 \leq y < 1 \\ \frac{1+\sqrt{y}}{3}, & 1 \leq y < 4 \\ 1, & 4 \leq y \end{cases}$$

- So

$$h(y) = \begin{cases} 0, & y < 0 \\ \frac{1}{3\sqrt{y}}, & 0 < y < 1 \\ \frac{1}{6\sqrt{y}}, & 1 < y < 4 \\ 0, & 4 < y \end{cases}$$

Transformations

- **Prop:** If g is strictly $\uparrow$ or strictly $\downarrow$, and $y = g(x)$ then $h(y) = f(x)/|g'(x)|$.

  [Proof: If g is $\uparrow$ then
  $H(y) = P(Y \le y) = P(g(X) \le g(x)) = P(X \le x) = F(x)$.
  Now apply the chain rule, and use that $\frac{dx}{dy} = 1/g'(x)$.
  Likewise if g is $\downarrow$, except now $H(y) = 1 - F(x)$.]
- **Prop:** If F is continuous and strictly $\uparrow$, and U is uniform on $[0,1]$, then $X \stackrel{d}{=} F^{-1}(U)$.

  [Proof. F^{-1} exists. And $P(F^{-1}(U) \le x) = P(F(F^{-1}(U)) \le F(x)) = P(U \le F(x)) = F(x)$]
- This is very useful computationally. A computer program (eg. Excel) will simulate random numbers $\sim F$ by first simulating uniform random numbers, and then applying F^{-1}. It doesn't need a separate simulation program for every possible F.