EC381/MN308
Probability and Some Statistics

Yannis Paschalidis

yannisp@bu.edu, http://ionia.bu.edu/

Dept. of Manufacturing Engineering

Dept. of Electrical and Computer Engineering

Center for Information and Systems Engineering

Lecture 9 - Outline

1. Functions of Random Variables
2. A Unified View of Discrete and Continuous Random Variables

3.5 Functions of Continuous RVs

A random variable X is transformed by a specified function $g(X)$ into a random variable:

$$Y = g(X).$$

What are the statistical properties of Y?

Types of Transformations:

- discrete → discrete
- continuous → discrete
- continuous → continuous
- continuous → mixed

In any of these cases, the expected value of Y, or of functions of Y, can be readily determined, but the PMF/ PDF are harder to determine

$$E[Y] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

$$E[Y^2] = E[g^2(X)]$$

$$\text{Var}[Y] = E[g^2(X)] - \{E[g(X)]\}^2$$

Discrete-to-Discrete Transformations

If X is a discrete RV taking on values $x_1, x_2, \ldots, x_n, \ldots$, then Y must also be a discrete RV with values $y_1 = g(x_1)$, $y_2 = g(x_2)$, … , $y_n = g(x_n)$, …

Some of these values may be the same, i.e., $g(x_i) = g(x_j)$ for some values of i and j and this will, of course, be reflected in the probabilities associated with Y

The PMF of Y is readily obtained via:

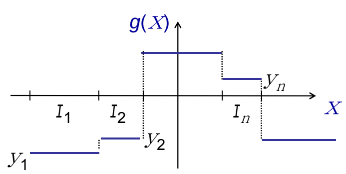
$$p_Y(y_j) = P\{Y = y_j\} = \sum p_X(x_i),$$

where the sum is over all i such that $g(x_i) = y_j$.

Continuous-to-Discrete Transformations

If the range of X can be partitioned into a (countable) set of intervals $I_1, I_2, \ldots I_n, \ldots$ such that $g(x)$ has a constant value y_j for all $x \in I_j$,

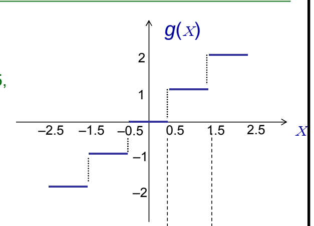
Then Y is a discrete RV taking values $y_1, y_2, \ldots, y_n, \ldots$ with probabilities $P\{X \in I_1\}$, $P\{X \in I_2\}$, … , $P\{X \in I_n\}$, …, respectively, as shown in the figure below:



Example: A quantizer function

$g(\bullet)$ is a staircase function

$g(x) = n$, for $n - 0.5 < x \leq n + 0.5$, n = integer



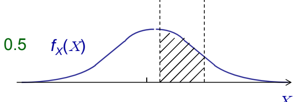
$P_Y(n) = P[Y = n]$

$= P[n - 0.5 < X \leq n + 0.5]$

= area under PDF $f_X(x)$

from $x = n - 0.5$ to $x = n + 0.5$

$= F_X(n + 0.5) - F_X(n - 0.5)$



Example: A quantizer function transforms an exponential Continuous RV into a Geometric Discrete RV

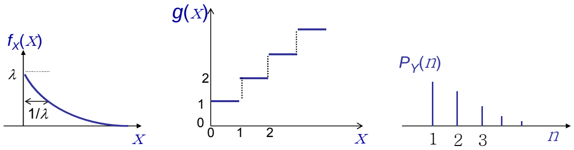
If X is an exponential RV with parameter λ and

$Y = \lceil X \rceil$ (smallest integer greater than or equal to x)

Then Y is a geometric random variable with parameter

$p = 1 - \exp(-\lambda)$ since $\int_0^1 \lambda \exp(-\lambda x)\,dx = 1 - \exp(-\lambda)$, so

$P_Y(n) = [1 - \exp(-\lambda)] \cdot [\exp(-\lambda)]^{n-1}$ for $n \geq 1$



Continuous-to-Continuous Transformations

If X is a continuous RV and $g(x)$ is a continuous function of x such that its derivative $g'(x)$ is not zero on any interval (this restriction simply avoids "flat spots" in $g(\bullet)$ that would give rise to mixed or discrete RVs, although allowing $g'(x) = 0$ for some values of x is acceptable),

Then, $Y = g(X)$ is a continuous RV.

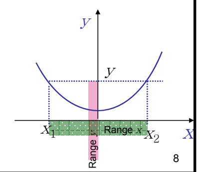
Procedure to determine the PDF of Y, given the PDF of X

1) Sketch the function $y = g(x)$ so that you can determine the range of values of X that correspond to any range of values of Y

2) Determine the CDF of Y in terms of the CDF of X:

$$F_Y(y) = P[-\infty < Y < y] = P[x_1 < X < x_2]$$
$$= F_X(x_2) - F_X(x_1)$$

3) Differentiate to get the PDF: $f_Y(y) = (d/dy)F_Y(y)$

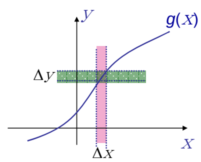


Special Case: $Y = g(X)$ is a continuous monotonic increasing function for the entire range of X. Given $f_X(x)$, determine $f_Y(y)$.

$P[y \leq Y \leq y + \Delta y] = P[x \leq X \leq x + \Delta x]$

$$f_Y(y)\Delta y = f_X(x)\Delta x$$

$$f_Y(y)\frac{dy}{dx} = f_X(x).$$



Similarly, for a monotonic decreasing function

$$-f_Y(y)\frac{dy}{dx} = f_X(x).$$

So, for any monotonic function: $f_Y(y)\left|\frac{dy}{dx}\right| = f_X(x)$

Example:

X is an exponential RV: $f_X(x) = \lambda \exp(-\lambda x), \quad x \geq 0$ and $Y = X^2, \quad X \geq 0$

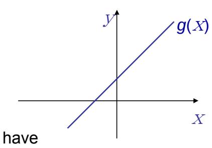
$$f_Y(y)\left|\frac{dy}{dx}\right| = f_X(x) \;\Rightarrow\; f_Y(y) \cdot 2x = f_X(x)$$

$$f_Y(y) = \frac{f_X(x)}{2x} = \frac{f_X\left(\sqrt{y}\right)}{2\sqrt{y}} = \frac{\lambda}{2\sqrt{y}}\exp\left(-\lambda\sqrt{y}\right)$$

Special Case: Linear Transformations

$Y = cX + d$, where $c > 0$

Given $f_X(x)$, determine $f_Y(y)$.



Since $g(X)$ is continuous and monotonic, we have

$$f_Y(y)\left|\frac{dy}{dx}\right| = f_X(x) \;\Rightarrow\; f_Y(y)|c| = f_X(x)$$

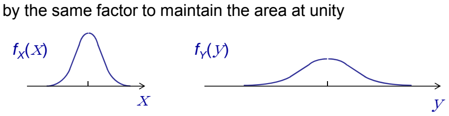
$$f_Y(y) = \frac{1}{|c|} f_X\left(\frac{y-d}{c}\right)$$

If $d = 0$, then, $f_Y(y) = \frac{1}{|c|} f_X\left(\frac{y}{c}\right)$

Thus, the function $f_Y(y)$ has the same shape as $f_X(x)$, but is stretched out by a factor of c (if $c > 1$) or squeezed down by a factor of c (if $c < 1$)

If the horizontal scale expands by a factor c, the vertical scale must compress by the same factor to maintain the area at unity

$c = 2$



If $d \neq 0$, then as before: $f_Y(y) = \frac{1}{|c|} f_X\left(\frac{y-d}{c}\right)$

i.e., $f_Y(y)$ is a scaled version of $f_X(x)$ that is translated rightward by d.

Example: Linear Transformation of a Gaussian RV

If X is a Gaussian RV, $N(\mu, \sigma^2)$, then $Y = cX + d$

is also a Gaussian RV, $N(c\mu + d, c^2\sigma^2)$.

Proof:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

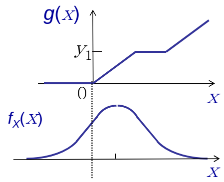
$$f_Y(y) = \frac{1}{|c|} f_X\left(\frac{y-d}{c}\right)$$

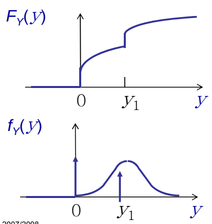
$$f_Y(y) = \frac{1}{|c|}\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{\left(\left(\frac{y-d}{c}\right)-\mu\right)^2}{2\sigma^2}\right] = \frac{1}{|c|\sigma\sqrt{2\pi}}\exp\left[-\frac{(y-d-\mu c)^2}{2c^2\sigma^2}\right]$$

$$= N(c\mu + d, c^2\sigma^2)$$

Continuous-to-Mixed Transformations

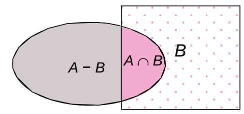
If X is a continuous RV and $g(x)$ is a continuous function of x with zero derivative on one or more intervals then $Y = g(X)$ is a mixed (or hybrid) RV





3.6 Conditional Probability for General RVs

Review: for random events



$$P[B|A] = \frac{P[A \cap B]}{P[A]} = \frac{P[A \cap B]}{P[A - B] + P[A \cap B]}$$

Review: for Discrete RVs:

$$P_{X|B}(x) = P(\{X = x\}|B) = \frac{P(\{X = x\} \cap B)}{P(B)}$$
$$= \frac{P(X = x, B)}{P(B)}$$

Special case: $B \subset S_X$ p. 83

$$P_{X|B}(x) = \begin{cases} \frac{P_X(x)}{P(B)} & x \in B \\ 0 & \text{otherwise} \end{cases}$$

$P_{X|B}(x)$ is a probability mass function with the usual properties:

$$P_{X|B}(x) \geq 0$$
$$\sum_{x \in S_X} P_{X|B}(x) = 1$$
$$\sum_{x \in B} P_{X|B}(x) = 1 \quad \text{if } B \subset S_X$$
$$\text{if } B \subset S_X, \text{ and } C \subset B, \ P[C|B] = \sum_{x \in C} P_{X|B}(x)$$

Conditional Probability for Arbitrary RVs

Conditional CDF: $F_{X|B}(a|B) \equiv P_{X \leq a|B}(x) = \frac{P(\{X \leq a\} \cap B)}{P(B)}$

Conditional PDF: $f_{X|B}(a|B) = \frac{d}{da} F_X(a|B) = \frac{\frac{d}{da} P(\{X \leq a\} \cap B)}{P(B)}$

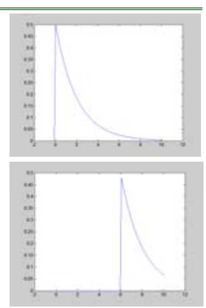
•Special case: $B \subset S_X$
–Truncate and rescale!

$$f_{X|B}(x) = \begin{cases} \frac{f_X(x)}{P(B)} & x \in B \\ 0 & \text{otherwise} \end{cases}$$

Example: Exponential RV truncated and rescaled

X is exponential, parameter $\lambda = 0.5$;

$B = \{X > 6\}$ (we know the RV > 6)



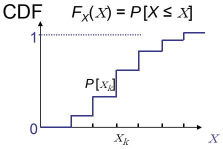
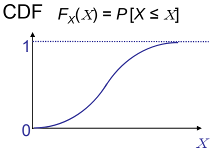
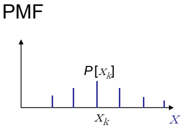
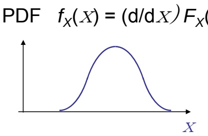
$$f_{X|B}(x) = \begin{cases} \frac{f_X(x)}{P(B)} & x \in B \\ 0 & \text{otherwise} \end{cases}$$
$$= \begin{cases} \frac{f_X(x)}{e^{-3}} & x \in B \\ 0 & \text{otherwise} \end{cases}$$
$$= \begin{cases} \frac{0.5e^{-0.5x}}{e^{-3}} & x \in B \\ 0 & \text{otherwise} \end{cases}$$

= e^{-3} since area · $f_X(x)$ from $x = 6$ to = $F_X(\infty) - F_X(6)$ = $_X(6) = e^{-0.5 \times 6} = e^{-3}$

For exponential RVs (and also Geometric RVs): residual life X_r is exponentially distributed, i.e.,

$P\{X_r > b \mid X > a\} = P\{X > b\}$, so that these RVs are "memoryless."

3.7 Unified View of Continuous & Discrete RVs

Discrete RV	Continuous RV
CDF $F_X(x) = P[X \leq x]$	CDF $F_X(x) = P[X \leq x]$
	
PMF	PDF $f_X(x) = (d/dx) F_X(x)$
	

Let's bring both of these under the same view.

PDF versus PMF

- Discrete RV defines a set of point masses on the axis, with total mass = 1
  - PMF $P_X(x)$ = mass at x = probability that x occurs
- Continuous RV defines a spread of the total unity probability mass along the axis
  - There is no probability mass at any point
- PDF of a continuous RV tells the density of the mass at each point
  - PDF is measured in units of probability mass/length
  - PDF is analogous to mass density, charge density, etc.

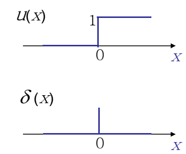
Define PDF functions for Discrete RVs using the concept of the impulse (delta) function

PDF = $f_X(x) = (d/dx) F_X(x)$ for both discrete & continuous RVs

Differentiate the "stairs" function using the unit step function $u(x)$ and the delta function $\delta(x)$:

$$u(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

$$\delta(x) = \frac{d}{dx}u(x) = \begin{cases} 0 & x \neq 0 \\ \infty & x = 0 \end{cases}$$



Properties of the Impulse Function

- The unit impulse is a "generalized function"
  - In principle it's not a proper function, as it has an infinite value
  - But, we can integrate against it
  - From a mathematician's perspective, it is a "distribution" or "generalized function"
- Given any well-behaved function $f(.)$ (e.g., bounded and measurable),

$$\int_{-\infty}^{\infty} f(x)\delta(x)dx = f(0)$$

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx = f(a)$$

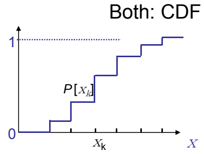
Unit impulse looks like a probability density!

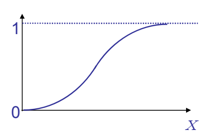
- It is nonnegative
- It integrates to 1: $\int_{-\infty}^{\infty} \delta(x)dx = 1$
- So, we will use it to describe PDFs for RVs with CDFs that are discontinuous
  - Discrete RVs
  - Mixed (or hybrid) RVs

Random Variable

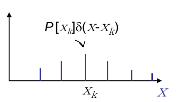
Discrete RV	Continuous RV

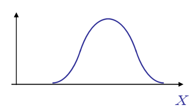
Both: CDF $F_X(x) = P[X \leq x]$





Both: PDF $f_X(x) = (d/dx) F_X(x)$





$f_X(x) = \sum_k P[x_k]\delta(x - x_k)$

Unified View of the Expectation of a Function of an RV

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$ for both continuous & discrete

If the RV is discrete: $f_X(x) = \sum_k P[x_k]\delta(x - x_k)$

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) \sum_k P[x_k]\delta(x - x_k)dx$$

$$= \sum_k P[x_k] \int_{-\infty}^{\infty} g(x)\delta(x - x_k)dx$$

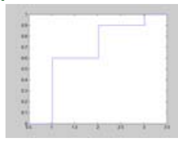
$$= \sum_k P[x_k] g(x_k)$$

i.e., $E[g(X)] = \sum_k g(x_k)P[x_k]$ as it should

Example

X is discrete RV, with $P[X = 1] = 0.6$, $P[X = 2] = 0.3$, $P[X = 3] = 0.1$

Determine the PDF and second moment $E[X^2]$.



CDF (use unit step function):

$$F_X(x) = 0.6u(x-1) + 0.3u(x-2) + 0.1u(x-3)$$

Differentiate to get PDF (use delta function):

$$f_X(x) = 0.6\delta(x-1) + 0.3\delta(x-2) + 0.1\delta(x-3)$$

Compute $E[X^2]$ using old way (PMF): $E[g(X)] = \sum_{n=1}^{3} g(n)P(X = n)$

$$= 0.6 \cdot 1 + 0.3 \cdot 4 + 0.1 \cdot 9 = 2.7$$

Compute via integral of PDF using new way:

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)dx$$
$$= \int_{-\infty}^{\infty} x^2(0.6\delta(x-1) + 0.3\delta(x-2) + 0.1\delta(x-3))dx$$
$$= 0.6 \cdot 1 + 0.3 \cdot 4 + 0.1 \cdot 9 = 2.7$$

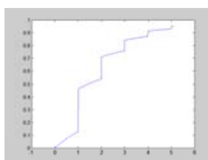
Mixed RVs

Example (admittedly somewhat artificial):

- Experiment:
  - 2 cashiers. Cashier 1 is "sloppy," cashier 2 is "meticulous." Meticulous one services customers exactly on the minutes.
  - Pick cashier 1 with probability 1/3 and cashier 2 with probability 2/3.
  - If cashier 1 is picked, generate delay *X* as an exponential RV with parameter $\lambda = 0.5$. If cashier 2 is picked, generate delay *X* as geometric RV with parameter $p = 0.5$.
- What is CDF of *X*?
  - Event *A*: cashier 1 picked.
  - Event *B*: cashier 2 picked.
  - Use Total Probability Theorem.

CDF

$$P(X \le a) = P(X \le a|A)P(A) + P(X \le a|B)P(B) \text{ Tot Prob Thm}$$
$$P(X \le a|A) = (1 - e^{-0.5a})u(a)$$
$$P(X \le a|B) = \sum_{1 \le n \le a} 0.5^n$$



PDF

$$f_X(a) = \frac{d}{da}[P(X \le a|A)P(A) + P(X \le a|B)P(B)]$$
$$= \frac{d}{da}[(1/3)(1 - e^{-0.5a})u(a) + (2/3)\sum_{1 \le n \le a} 0.5^n]$$
$$= (1/6)e^{-0.5a}u(a) + (2/3)\sum_{n=1}^{\infty}(0.5)^n\delta(a - n)$$