The pdf of a continuous RV

- The pdf $f_X(u)$ of a continuous random variable is the derivative of its CDF
- The pdf has two properties
  - $f_X(u) \geq 0$
  - Total area under the pdf curve $f_X(u)$ from $-\infty$ to ∞ is 1

Fancy statement: $\int_{-\infty}^{\infty} f_X(u)\,du = 1$

Other properties enjoyed by the pdf

- The pdf is a nonnegative function that has unit area between the pdf curve and the horizontal axis
- $f_X(+\infty) = \lim_{u \to +\infty} f_X(u) = 0$
- $f_X(-\infty) = \lim_{u \to -\infty} f_X(u) = 0$
- Compare to $F_X(+\infty) = 1$ and $F_X(-\infty) = 0$

Probabilities from the pdf

- $P\{a \leq \mathbf{X} \leq a + \delta a\} \approx f_X(a) \cdot \delta a$
- $P\{a < \mathbf{X} < b\}$ = Area under the pdf curve between a and b $= \int_a^b f_X(u)\,du$
- $P\{\mathbf{X} = u\} = 0$ for all real numbers u
- $P\{a < \mathbf{X} < b\}$, $P\{a < \mathbf{X} \leq b\}$, $P\{a \leq \mathbf{X} < b\}$, and $P\{a \leq \mathbf{X} \leq b\}$ all have the same value

Relationship between CDF and pdf

- The value of the CDF at the point u = 5 is $F_X(5)$ = area under pdf $f_X(u)$ from $-\infty$ to 5
- The CDF is not the antiderivative (or indefinite integral) of the pdf; it is the definite integral
- Re-read your calculus books to refresh your understanding of the Fundamental Theorem of Calculus

ECE 313 Survival Guide

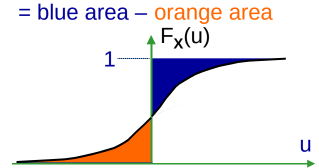
- Always, always, always sketch the pdf (or CDF) curve before you do anything else
- Indicate the desired probability as an area (e.g. by shading) on the sketch
- If you use an integral to find the area, set the limits with the help of the sketch
- Do not use indefinite integrals
- All integrals must have limits

Expectation of an arbitrary RV

- The expected value $E[\mathbf{X}]$ of an arbitrary random variable $\mathbf{X}$ can be defined as

$$E[\mathbf{X}] = \int_0^{\infty} [1 - F_X(u)]\,du - \int_{-\infty}^{0} F_X(u)\,du$$

= blue area – orange area



Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign

Expectation of a continuous RV

- $E[\mathbf{X}] = \int_{-\infty}^{\infty} u \cdot f_X(u)\,du$
  is a much more useful formula for computational purposes than
  $E[\mathbf{X}] = \int_{0}^{\infty} [1 - F_X(u)]\,du - \int_{-\infty}^{0} F_X(u)\,du$
- Since an integral is a glorified sum, the expectation integral is the analogue of the result $E[\mathbf{X}] = \sum u_i \cdot p(u_i)$ for discrete RVs

Who's on first?

- The formulas
  $E[\mathbf{X}] = \int_{-\infty}^{\infty} u \cdot f_X(u)\,du$ and $E[\mathbf{X}] = \sum u_i \cdot p(u_i)$
  were "discovered" first and used widely
- Only later was it realized that both were special instances that could be derived from the more general definition
  $E[\mathbf{X}] = \int_{0}^{\infty} [1 - F_X(u)]\,du - \int_{-\infty}^{0} F_X(u)\,du$

The same thing, only different…

- $E[\mathbf{X}] = \int_{-\infty}^{\infty} u \cdot f_X(u)\,du$
  is analogous to $E[\mathbf{X}] = \sum u_i \cdot p(u_i)$ for discrete random variables
- $P\{u \le \mathbf{X} \le u + \delta u\} \approx f_X(u) \cdot \delta u$
- Probability mass of approximately $f_X(u) \cdot \delta u$ is at distance u from the origin
- It contributes a moment of $\approx u \cdot f_X(u) \cdot \delta u$
- $E[\mathbf{X}]$ = total moment of probability mass

The same interpretations hold

- $E[\mathbf{X}] = \int_{-\infty}^{\infty} u \cdot f_X(u)\,du = \mu$ or μ_X
- Interpretations of $E[\mathbf{X}]$
  - Average value of **X** over many trials
  - Moment about origin of prob. masses
  - μ is the location of the center of mass
  - μ is the fair price to play a game in which the winnings on each trial are **X**

$\infty - \infty$ is not allowed

- $u \cdot f_X(u) \ge 0$ for $u > 0$
- $u \cdot f_X(u) \le 0$ for $u < 0$
- Break integral for $E[\mathbf{X}]$ into two integrals, over $(-\infty, 0)$ and $(0, \infty)$ respectively
- Both integrals should not be infinite
- Cauchy RV has pdf $[(\pi(1+u^2)]^{-1}$
- The pdf is symmetric about the origin
- Center of mass = 0? No. The integral of $u \cdot [(\pi(1+u^2)]^{-1}$ is of the form $\infty - \infty$
- $E[\mathbf{X}]$ is undefined for the Cauchy RV

LOTUS for continuous RVs

- We can transform **X** into **Y** by applying function g to the observed value of **X**
- LOTUS: If **Y** = g(**X**) where **X** is a continuous random variable and g(•) is a measurable function, then
  $E[\mathbf{Y}] = E[g(\mathbf{X})] = \int_{-\infty}^{\infty} g(u) \cdot f_X(u)\,du$
- Slide 39 in Lecture 23 mistakenly asserts that g(•) has to be a continuous function.
- Not so!!

What's measurable mean, anyway?

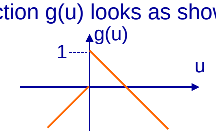
- Measurable function is a technical term requiring mathematical sophistication far beyond what is expected in this course
- The requirement that g(•) be measurable is needed to ensure that the CDF of **Y** is properly defined
- Good news: All the functions that we will encounter in ECE 313 are measurable
- What? Me worry? — Alfred E. Neumann

Correction of error from Lecture 23

- In Lecture 23, it was mistakenly asserted that LOTUS could not be used with a staircase function such as a quantizer
- Not so; LOTUS applies to staircase functions, or discontinuous functions, or …
- All that happens is that the range of integration is broken up into pieces
- This illustrates the danger in putting one's mouth in gear while brain is disengaged

Example

- Let **Y** = **X** for **X** ≤ 0, and **Y** = 1 – **X** for **X** > 0
- The function g(u) looks as shown



- $E[\mathbf{Y}] = E[g(\mathbf{X})] = \int_{-\infty}^{\infty} g(u) \cdot f_{\mathbf{X}}(u)\,du$

$$= \int_{-\infty}^{0} u \cdot f_{\mathbf{X}}(u)\,du + \int_{0}^{\infty} (1-u) \cdot f_{\mathbf{X}}(u)\,du$$

Using LOTUS

- LOTUS requires the computation of a definite integral over the entire real line
- Limits are generally not a problem
- However, sketching g(u) and f(u) on the same axes (and possibly g(u)•f(u)) can often help
- Example: If g(u) > 0 for all u but your computations result in E[g(**X**)] < 0, then you have made a mistake

Applications of LOTUS

- **Y** = a**X** + b, where a and b are constants
- E[a**X** + b] = a•E[**X**] + b
- Expectation is a linear operation: the expectation of a sum is the sum of the expectations
- E[a**X**] = a•E[**X**]
- E[b] = b
- Masses in the pdf of **Y** = a**X** are "further away" from the origin by a factor of a, and so is the center of mass "further away"

This will hurt for just a few moments …

- E[**X** – a] is the (first) moment of **X** about a
- $E[(\mathbf{X} - a)^n]$ = n-th moment of **X** about a
- $E[\mathbf{X}^n]$ is called the n-th moment of **X**
- $E[(\mathbf{X} - \mu)^n]$ is the n-th central moment of **X**
- E[**X** – μ], the first central moment of **X** is 0
- $E[(\mathbf{X} - \mu)^2]$, the second central moment of **X**, is commonly called the variance of **X**

Variance

- E[(**X** – μ)^2], the second central moment of **X**, called the variance of **X**
- The variance of **X** is denoted by var(**X**) or σ^2 or $\sigma_X{}^2$. var(**X**) > 0 for continuous RVs
- σ^2 is the moment of inertia about μ; σ can be thought of as the radius of gyration
- $\sigma^2 = E[X^2] - \mu^2 = E[X^2] - (E[X])^2$
  $= E[X^2] - E^2[X]$

The variance measures the spread

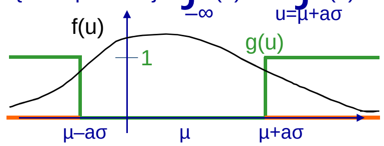
- $E[aX] = a \bullet E[X] = a \bullet \mu_X$
- $E[(aX)^2] = a^2 \bullet E[X^2]$
- $\text{var}(aX) = E[(aX)^2] - (E[aX])^2$
  $= a^2 \bullet E[X^2] - (a \bullet \mu_X)^2$
  $= a^2 \bullet (E[X^2] - (\mu_X)^2)$
  $\text{var}(aX) = a^2 \bullet \text{var}(X)$
- $\text{var}(aX + b) = \text{var}(aX) = a^2 \bullet \text{var}(X)$

The Chebyshev Inequality

- Chebyshev inequality: For any random variable with mean μ and variance σ^2
  P{½**X** – μ½≥ aσ} ≤ $1/a^2$
- No more than $1/a^2$ of the probability mass is at distance of aσ or more from the mean
- Usually, far less than $1/a^2$ mass is so far away from the mean

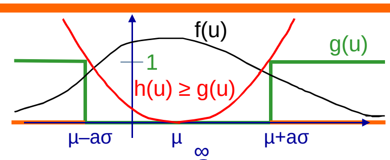
Proof of Chebyshev Inequality I

- P{½**X** – μ½≥ aσ} = $\int_{-\infty}^{u=\mu-a\sigma} f(u)\,du + \int_{u=\mu+a\sigma}^{\infty} f(u)\,du$



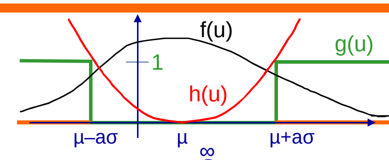
- P{½**X** – μ½≥ aσ} = $\int_{-\infty}^{\infty} g(u)f(u)\,du$

Proof of Chebyshev Inequality II



- P{½**X** – μ½≥ aσ} = $\int_{-\infty}^{\infty} g(u)f(u)\,du$
- P{½**X** – μ½≥ aσ} ≤ $\int_{-\infty}^{\infty} h(u)f(u)\,du$

Proof of Chebyshev Inequality III



- P{½**X** – μ½≥ aσ} ≤ $\int_{-\infty}^{\infty} h(u)f(u)\,du$
- Choose $h(u) = (u-\mu)^2/(a\sigma)^2$
- The integral is $\text{var}(X)/(a\sigma)^2 = 1/a^2$

Uniform random variables

- **X** is said to be a uniform random variable on an interval (or uniformly distributed over an interval) if its pdf is
  - constant over the interval
  - 0 outside the interval (i.e. otherwise)
- On the interval, the pdf has constant value

  1/{length of interval}
- Why?

Basic information re uniform RV

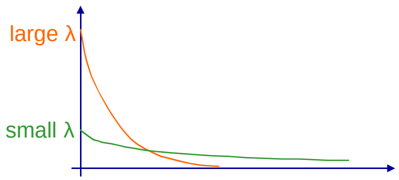
- **X** is uniformly distributed over [a, b]
- Its pdf has value 1/(b – a) for $a \le u \le b$
- E[**X**] = (a + b)/2 = midpoint of interval
- var(**X**) = $(b - a)^2/12 = (\text{length})^2/12$
- Uniform random variables are used to model complete lack of knowledge, or complete indifference
- Principle of insufficient reason…

Exponential random variables

- Exponential random variables arise in studies of waiting times, service times, etc
- **X** is called an exponential random variable with parameter λ if its pdf is given by
  - $f(u) = \lambda \cdot \exp(-\lambda u)$ for $u \ge 0$
  - $f(u) = 0$ for $u < 0$
- Note that $\lambda > 0$. If λ is large, the pdf is large at 0 but also decays away rapidly

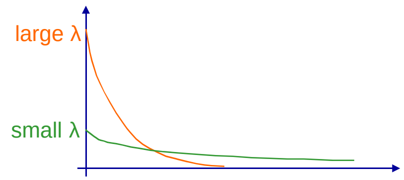
Comparison of pdfs

- If λ is large, the pdf has large value at 0 but also decays away rapidly
- Two pdfs must always cross. Why?



Mean & variance of exponential RV

- λ is called the scale parameter
- E[**X**] = 1/λ var(**X**) = $1/\lambda^2$
- If λ is large, most of the mass is close to 0



Gamma random variables

- Exponential random variables are a subclass of gamma random variables
- **X** is called a gamma random variable with parameters (t, λ) if its pdf is given by
  - $f(u) = \lambda \cdot \exp(-\lambda u) \cdot (\lambda u)^{t-1}/\Gamma(t)$ for $u > 0$
  - $f(u) = 0$ for $u \le 0$
- $t > 0$, $\lambda > 0$; Γ(t) is the gamma function
- Γ(t) is just a number in the formula for f(u)

All you wanted to know about Γ(t)…

- The value of Γ(t), the gamma function, can be given in terms of an integral
- $\Gamma(t) = \int_0^\infty x^{t-1} \bullet \exp(-x)\, dx, \quad t > 0$
- Γ(1) = 1
- For t > 1, integration by parts shows that Γ(t) = (t–1)•Γ(t–1) = (t–1)•(t–2)•Γ(t–2) = …
- If t is an integer, Γ(t) = (t – 1)!
- For non-integer t, Γ(t) = (t–1)•…•Γ(t – ëtû)

… but were afraid to ask!

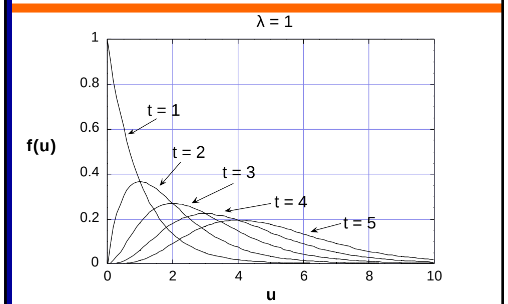
- For non-integer t, Γ(t) = (t–1)•…•Γ(t – ëtû)
- ëtû is the integer part of t
- t – ëtû is the fractional part of t: 0 < t–ëtû < 1
- For 0 < t < 1, the value of Γ(t) must be computed via numerical integration
- For 0 < t < 1, Γ(t) is tabulated in various reference books?
- Does your calculator do gamma functions?
- $\Gamma(1/2) = \sqrt{\pi}$

Scale and order parameters

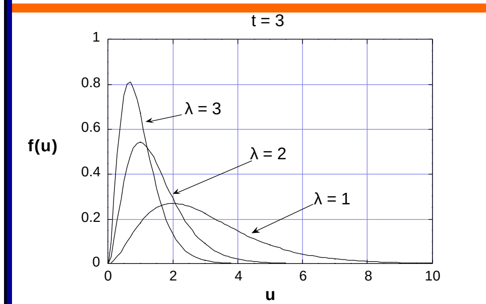
- **X** is called a gamma random variable with parameters (t, λ) if its pdf is given by

  $f(u) = \lambda \bullet \exp(-\lambda u) \bullet (\lambda u)^{t-1}/\Gamma(t)$ for $u > 0$
- Area under $\lambda \bullet \exp(-\lambda u) \bullet (\lambda u)^{t-1}$ from 0 to ∞ is Γ(t). Show this by a change of variables
- λ is called the scale parameter
- t is called the order parameter
- A (1, λ) gamma RV is an exponential RV

Gamma pdfs: different values of t



Gamma pdfs: different values of λ



Mean & variance of gamma RVs

- If **X** is a gamma random variable with parameters (t, λ), then

  $E[\mathbf{X}] = t/\lambda$ and $\mathrm{var}(\mathbf{X}) = t/\lambda^2$
- If λ is large, most of the mass is close to 0
- If t is large, most of the mass is far away from 0
- The CDF of a gamma RV is related to the incomplete gamma function

That which we call a rose…

- Gamma RVs arise in lots of applications
- In queueing theory (ECE 338, ECE 467), a gamma RV with parameters (n, λ) is called an n-Erlang random variable
- In statistical applications, a gamma RV with parameters (n/2, 1/2) is called a chi-square (or χ^2 random variable with n degrees of freedom

…by any other name would …

- The exponential RV with parameter λ is analogous to the geometric RV with parameter p — both denote the waiting time till some event occurs
- E[**X**] = 1/λ ↔ E[**X**] = 1/p
- The gamma RV with parameters (r, λ) is analogous to the negative binomial RV with parameters (r, p) — they denote the waiting time for the r-th occurrence
- E[**X**] = r/λ ↔ E[**X**] = r/p

… smell as sweet

- In the next lecture, we shall make these analogies more precise and relate all these random variables (as well as Poisson random variables) to a very important idea:

The Poisson Process

Summary

- We have revisited LOTUS (correctly!) and looked at its consequences
- We have studied uniform RVs
- We have defined the exponential and gamma RVs, and studied their basic properties (and analogies to some previously studied discrete RVs)
- We have threatened that we will be studying the Poisson process next