1. Course Introduction + Bootcamps

Dave Goldsman

H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

3/2/20

Georgia
Tech

Outline of Lessons

Georgia
Tech

Lesson 1.1 — Introduction to Probability and Statistics

Next Few Lessons:

- This Probability and Statistics Intro
- Set Theory Bootcamp
- Calculus Bootcamps

Mathematical Models for describing observable phenomena:

- Deterministic
- Probabilistic

Deterministic Models

- Ohm's Law ($I = E/R$) (There's no place like Ohm.)
- Drop an object from height h_0. After t seconds, height $h(t) = h_0 - 16t^2$.
- Deposit \$1000 in a checking account, continuously compounding at 3%. At time t, it's worth \$$1000e^{0.03t}$.

Georgia Tech

Probabilistic Models — Involve uncertainty

- How much snow will fall tomorrow?
- Will IBM make a profit this year?
- Should I buy a call or put option?
- Can I win in blackjack if I use a certain strategy?
- What is the cost-effectiveness of a new drug?
- Which horse will win the Kentucky Derby?

Some Cool Examples

- Birthday Problem — Assume all 365 days have equal probability of being a person's birthday (ignore freaks born on Feb. 29). Then...

  If there are **23** students in a class, the odds are better than 50–50 that there will be a match.

  If there are **50** students, the probability is about 97%!

- Monopoly — In the long run, the property having the highest probability of being landed on is Illinois Ave.

- Stock Market — Monkeys randomly selecting stocks could have outperformed most market analysts during the past year.

Georgia
Tech

- Poker — Pick 5 cards from a standard deck. Then

  $P(\text{exactly 2 pairs}) \approx 0.0475$,

  $P(\text{full house}) \approx 0.00144$,

  $P(\text{flush}) \approx 0.00198$.

- A couple has two kids and at least one is a boy. What's the probability that BOTH are boys?

  Possibilities: GG, BG, GB, BB. Eliminate GG since we know that there's at least one boy. Then $P(\text{BB}) = 1/3$.

- Ask Marilyn. You are a contestant at a game show. Behind one of three doors is a car; behind the other two are goats. You pick door A. Monty Hall opens door B and reveals a goat. Monty offers you a chance to switch to door C. What should you do? Answer: SWITCH!

Georgia
Tech

- Vietnam Draft Lottery — not as "fair" as you might think!

- Which is the most popular soft drink? Well, in Atlanta, we know the answer to that one!

- Why are some election polls so incredibly wrong?

- How do they do real-time updates of win probabilities as a basketball game progresses?

- How can you simulate randomness on a computer, and what can you use it for?

- How can you tell if the quality of an item that your manufacturing plant is producing has started to get worse?

Georgia
Tech

Working Definitions

Probability — Methodology that describes the random variation in systems. (We'll spend about 50% of our time on this.)

Statistics — Uses sample data to draw general conclusions about the population from which the sample was taken. (50% of our time.)

Georgia
Tech

Outline of Lessons

Georgia
Tech

Lesson 1.2 — The Joy of Sets Bootcamp

Definition: A **set** is a collection of objects. Members of a set are called **elements**.

Notation:

$A, B, C, \ldots$ for sets

$a, b, c, \ldots$ for elements

$\in$ for membership, e.g., $x \in A$

$\notin$ for non-membership, e.g., $x \notin A$

U is the **universal set** (i.e., everything)

$\emptyset$ is the **empty set**.

Georgia
Tech

Examples:

$A = \{1, 2, \ldots, 10\}$. $2 \in A$; $49 \notin A$.

$B = \{\text{basketball, baseball}\}$

$C = \{x \mid 0 \le x \le 1\}$ ("$\mid$" means "such that")

$D = \{x \mid x^2 = 9\} = \{\pm 3\}$ (either is fine)

$E = \{x \mid x \in \mathbb{R}, x^2 = -1\} = \emptyset$ ($\mathbb{R}$ is the real line)

Definition: If every element of set A is an element of set B then A is a **subset** of B, i.e., $A \subseteq B$.

Definition: $A = B$ iff (if and only if) $A \subseteq B$ and $B \subseteq A$.

Properties:
$\emptyset \subseteq A$; $A \subseteq U$; $A \subseteq A$
$(A \subseteq B$ and $B \subseteq C) \Rightarrow$ (implies) $A \subseteq C$

Remark: The order in which the elements of a set are listed is immaterial, e.g., $\{a, b, c\} = \{b, c, a\}$.

Georgia
Tech

Definitions: The **complement** of A with respect to U is
$\bar{A} \equiv \{x \mid x \in U \text{ and } x \notin A\}$.

Remark: Don't confuse complement with compliment! ("You are one fine-lookin' set!")

The **intersection** of A and B is $A \cap B \equiv \{x \mid x \in A \text{ and } x \in B\}$.

The **union** of A and B is $A \cup B \equiv \{x \mid x \in A \text{ or } x \in B \text{ (or both)}\}$.

(Remember **Venn diagrams**?)

Georgia
Tech

Example:

Suppose $U = \{\text{letters of the alphabet}\}$, $A = \{\text{vowels}\}^{\star}$, and $B = \{a, b, c\}$.

*We'll ignore the fact that y and w are sometimes vowels (believe it or not)!
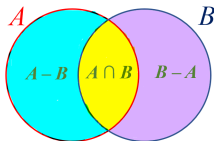
Then

$\bar{A} = \{\text{consonants}\}$
$A \cap B = \{a\}$
$A \cup B = \{a, b, c, e, i, o, u\}.$ $\square$

Definition: If $A \cap B = \emptyset$, then A and B are **disjoint** (or **mutually exclusive**).

More Definitions:

Minus: $A - B \equiv A \cap \bar{B}$

Symmetric difference or **XOR**:

$$A \,\Delta\, B \;\equiv\; (A - B) \cup (B - A) \;=\; (A \cup B) - (A \cap B)$$

Georgia
Tech

The **cardinality** of A, denoted by $|A|$, is the number of elements in A. A is **finite** if $|A| < \infty$.

Examples:

$A = \{3, 4\}$ is finite, since $|A| = 2$.

$B = \{1, 2, 3, \ldots\}$ is **countably infinite**, i.e., $|B| = \aleph_0$ (look up this symbol)!

$C = \{x \mid x \in [0, 1]\}$ is **uncountably infinite**, i.e., $|C| = \aleph_1$ (look it up)!

Georgia
Tech

Laws of Operation:

1. **Complement Law:** $A \cup \bar{A} = U$, $A \cap \bar{A} = \emptyset$, $\bar{\bar{A}} = A$

2. **Commutative:** $A \cup B = B \cup A$, $A \cap B = B \cap A$

3. **Associative:** $A \cup (B \cup C) = (A \cup B) \cup C$,
   $A \cap (B \cap C) = (A \cap B) \cap C$

4. **Distributive:** $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$,
   $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

5. **DeMorgan's:** $\overline{A \cup B} = \bar{A} \cap \bar{B}$, $\overline{A \cap B} = \bar{A} \cup \bar{B}$

Proofs: Easy. Could use Venn diagrams or many other ways.

Georgia
Tech

Outline of Lessons

Georgia
Tech

Lesson 1.3 — Calculus Bootcamp: Introduction + Derivatives

Definition: The **function** $f(x)$ maps values of x from a certain **domain** X to a certain **range** Y, which we can denote by the shorthand $f : X \rightarrow Y$.

Example: If $f(x) = x^2$, then the function takes x-values from the real line $\mathbb{R}$ to the nonnegative portion of the real line $\mathbb{R}^+$.

Definition: We say that $f(x)$ is a **continuous** function if, for any x_0 and $x \in X$, we have $\lim_{x \to x_0} f(x) = f(x_0)$, where "lim" denotes a **limit** and $f(x)$ is assumed to exist for all $x \in X$.

Example: The function $f(x) = 3x^2$ is continuous for all x. The function $f(x) = \lfloor x \rfloor$ (round down to the nearest integer, e.g., $\lfloor 3.4 \rfloor = 3$) has a "jump" discontinuity at any integer x. $\square$

Georgia Tech

Definition: The **inverse** of a function $f : X \to Y$ is (informally) the "reverse" mapping $g : Y \to X$ such that $f(x) = y$ if and only if $g(y) = x$ for all appropriate x and y. The inverse of f is often written as f^{-1}, and is especially useful if $f(x)$ is a strictly increasing or strictly decreasing function. Note that $f^{-1}(f(x)) = x$.

Examples: If $f(x) = x^3$, then we have $f^{-1}(y) = y^{1/3}$. If $h(x) = e^x$, then $h^{-1}(y) = \ell\text{n}(y)$. $\quad \square$

Definition: If $f(x)$ is continuous, then it is **differentiable** (has a **derivative**) if

$$\frac{d}{dx} f(x) \; \equiv \; f'(x) \; \equiv \; \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}$$

exists and is well-defined for any given x. Think of the derivative as the slope of the function.

Georgia
Tech

Examples: Some well-known derivatives are:

$$[x^k]' = kx^{k-1},$$

$$[c^x]' = c^x \, \ell \mathrm{n}(c), \ \text{ and so } \ [e^x]' = e^x \ (\text{since } \ell \mathrm{n}(e) = 1),$$

$$[\sin(x)]' = \cos(x),$$

$$[\cos(x)]' = -\sin(x),$$

$$[\ell \mathrm{n}(x)]' = \frac{1}{x},$$

$$[\arctan(x)]' = \frac{1}{1+x^2}. \quad \Box$$

Georgia
Tech

Theorem: Some well-known properties of derivatives are:

$$[af(x) + b]' = af'(x),$$

$$[f(x) + g(x)]' = f'(x) + g'(x),$$

$$[f(x)g(x)]' = f'(x)g(x) + f(x)g'(x) \quad \text{(product rule)},$$

$$\left[\frac{f(x)}{g(x)}\right]' = \frac{g(x)f'(x) - f(x)g'(x)}{g^2(x)} \quad \text{(quotient rule)}^1,$$

$$[f(g(x))]' = f'(g(x))g'(x) \quad \text{(chain rule)}^2.$$

---

[1] Ho dee Hi minus Hi dee Ho over Ho Ho.

[2] www.youtube.com/watch?v=gGAiW5dOnKo

Georgia
Tech

Example: Suppose that $f(x) = x^2$ and $g(x) = \ell n(x)$. Then

$$
\begin{aligned}
[f(x)g(x)]' &= f'(x)g(x) + f(x)g'(x) \\
&= 2x\ell n(x) + x(1/x) \\
&= 2x\ell n(x) + x,
\end{aligned}
$$

$$
\begin{aligned}
\left[\frac{f(x)}{g(x)}\right]' &= \frac{g(x)f'(x) - f(x)g'(x)}{g^2(x)} \\
&= \frac{[\ell n(x)]2x - x^2(1/x)}{\ell n^2(x)} \\
&= \frac{2x\ell n(x) - x}{\ell n^2(x)},
\end{aligned}
$$

$$
[f(g(x))]' = \left[(g(x))^2\right]' = 2g(x)g'(x) = \frac{2\ell n(x)}{x}. \quad \square
$$

Georgia
Tech

Remark: The second derivative $f''(x) \equiv \frac{d}{dx}f'(x)$ and is the "slope of the slope." If $f(x)$ is "position," then $f'(x)$ can be regarded as "velocity," and $f''(x)$ as "acceleration."

The minimum or maximum of $f(x)$ can only occur when the slope of $f(x)$ is zero, i.e., only when $f'(x) = 0$, say at the **critical point** $x = x_0$. Exception: Check the endpoints of your interval of interest as well.

Then if $f''(x_0) < 0$, you get a max; if $f''(x_0) > 0$, you get a min; and if $f''(x_0) = 0$, you get a **point of inflection**.

Example: Find the value of x that minimizes $f(x) = e^{2x} + e^{-x}$. The minimum can only occur when $f'(x) = 2e^{2x} - e^{-x} = 0$. After a little algebra, we find that this occurs at $x_0 = -(1/3)\ell n(2) \approx -0.231$. It's also easy to show that $f''(x) > 0$ for all x; and so x_0 yields a minimum. $\square$

Georgia
Tech

Outline of Lessons

Georgia
Tech

Lesson 1.4 — Calculus Bootcamp: Integration and Beyond

Integration

Definition: The function $F(x)$ having derivative $f(x)$ is called the **antiderivative** (or **indefinite integral**). It is denoted by $F(x) = \int f(x)\, dx$.

Fundamental Theorem of Calculus: If $f(x)$ is continuous, then the area under the curve for $x \in [a, b]$ is denoted and given by the **definite integral** [3]

$$\int_a^b f(x)\, dx \; \equiv \; F(x)\Big|_a^b \; \equiv \; F(b) - F(a).$$

---

[3]"I'm *really* an integral!"

Georgia
Tech

Example: Some well-known indefinite integrals are:

$$\int x^k \, dx \; = \; \frac{x^{k+1}}{k+1} + C \;\; \text{for } k \neq -1,$$

where C is an arbitrary constant,

$$\int \frac{dx}{x} \; = \; \ell\text{n}|x| + C,$$

$$\int e^x \, dx \; = \; e^x + C,$$

$$\int \cos(x) \, dx \; = \; \sin(x) + C,$$

$$\int \frac{dx}{1 + x^2} \; = \; \arctan(x) + C. \quad \square$$

Georgia
Tech

Example: It is easy to see that

$$
\begin{aligned}
\int \frac{d\,\mathrm{cabin}}{\mathrm{cabin}} &= \ell\mathrm{n}|\mathrm{cabin}| + C \quad \odot \\
&= \text{houseboat.} \quad \odot
\end{aligned}
$$

Theorem: Some well-known properties of definite integrals are:

$$
\int_a^a f(x)\,dx = 0,
$$

$$
\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx,
$$

$$
\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.
$$

Georgia
Tech

Theorem: Some other properties of general integrals are:

$$\int [f(x) + g(x)] \, dx \;=\; \int f(x) \, dx + \int g(x) \, dx,$$

$$\int f(x)g'(x) \, dx \;=\; f(x)g(x) - \int g(x)f'(x) \, dx \quad \text{(integration by parts)}[4],$$

$$\int f(g(x))g'(x) \, dx \;=\; \int f(u) \, du \quad \text{(substitution rule with } u = g(x))[5].$$

---

[4] www.youtube.com/watch?v=OTzLVIc-O5E
[5] www.youtube.com/watch?v=eswQl-hcvU0

Georgia
Tech

Example: To demonstrate integration by parts on a definite integral, let $f(x) = x$ and $g'(x) = e^{2x}$, so that $g(x) = e^{2x}/2$. Then

$$
\begin{aligned}
\int_0^1 xe^{2x}\, dx &= \int_0^1 f(x)g'(x)\, dx \\[2mm]
&= f(x)g(x)\Big|_0^1 - \int_0^1 g(x)f'(x)\, dx \quad \text{(parts)} \\[2mm]
&= \frac{xe^{2x}}{2}\Big|_0^1 - \int_0^1 \frac{e^{2x}}{2}\, dx \\[2mm]
&= \frac{e^2}{2} - \frac{e^{2x}}{4}\Big|_0^1 \\[2mm]
&= \frac{e^2 + 1}{4}. \quad \square
\end{aligned}
$$

Georgia
Tech

Definition: Derivatives of arbitrary order k can be written as $f^{(k)}(x)$ or $\frac{d^k}{dx^k} f(x)$. By convention, $f^{(0)}(x) = f(x)$.

The **Taylor series expansion** of $f(x)$ about a point a is given by

$$f(x) \ = \ \sum_{k=0}^{\infty} \frac{f^{(k)}(a)(x-a)^k}{k!}.$$

The **Maclaurin series** is simply Taylor expanded around $a = 0$.

Georgia Tech

Example: Here are some famous Maclaurin series:

$$\sin(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!},$$

$$\cos(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!},$$

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Example: And while we're at it, here are some miscellaneous sums that you should know:

$$\sum_{k=1}^{n} k \;=\; \frac{n(n+1)}{2},$$

$$\sum_{k=1}^{n} k^2 \;=\; \frac{n(n+1)(2n+1)}{6},$$

$$\sum_{k=0}^{\infty} p^k \;=\; \frac{1}{1-p} \;\; (\text{for } -1 < p < 1).$$

Theorem: Occasionally, we run into trouble when taking indeterminate ratios of the form $0/0$ or ∞/∞. In such cases, **L'Hôspital's Rule**[6] is useful: If the limits $\lim_{x \to a} f(x)$ and $\lim_{x \to a} g(x)$ both go to 0 or both go to ∞, then

$$\lim_{x \to a} \frac{f(x)}{g(x)} \;=\; \lim_{x \to a} \frac{f'(x)}{g'(x)}.$$

Example: L'Hôspital shows that

$$\lim_{x \to 0} \frac{\sin(x)}{x} \;=\; \lim_{x \to 0} \frac{\cos(x)}{1} \;=\; 1. \quad \square$$

Georgia
Tech

---

[6]This rule makes me sick.

Double Integration

We'll have occasion to calculate several double integrals. Whereas our usual (single) integrals get us the area under a curve, double integrals represent the *volume* under a three-dimensional function.

Example: The volume under $f(x, y) = 8xy$ over the region $0 < x < y < 1$ is given by

$$\int_0^1 \int_0^y f(x, y)\, dx\, dy = \int_0^1 \int_0^y 8xy\, dx\, dy = \int_0^1 4y^3\, dy = 1. \quad \square$$

We can usually swap the order of integration to get the same answer:

$$\int_0^1 \int_x^1 8xy\, dy\, dx = \int_0^1 4x(1 - x^2)\, dx = 1. \quad \square$$

Georgia
Tech

Saved by Zero! (How to Solve for a Root)

Suppose that we want to solve some equation $g(r) = 0$ for a root $r^\star$, where $g(r)$ is a nicely behaved continuous function.

We can use:

- trial-and-error or some sort of linear search — that's for losers! ☹
- **bisection method** — pretty fast! ☺
- **Newton's method** — really fast! 😜

Intermediate Value Theorem (IVT): If $g(\ell)g(u) < 0$, then there is a zero $r^\star \in [\ell, u]$. In other words, if (i) $g(\ell) < 0$ and $g(u) > 0$ or (ii) $g(\ell) > 0$ and $g(u) < 0$, then $g(r)$ crosses 0 somewhere between ℓ and u.

Georgia
Tech

Bisection uses the IVT to hone in on a zero via sequential bisecting:

- Initialization: Find lower and upper bounds $\ell_0 < u_0$ such that $g(\ell_0)g(u_0) < 0$. Then the IVT implies that $r^\star \in [\ell_0, u_0]$.

- For $j = 0, 1, 2, \ldots$,
  - Let the midpoint of the current interval be $r_j \leftarrow (\ell_j + u_j)/2$.
  - If $g(r_j)$ is sufficiently close to 0, or the interval width $u_j - \ell_j$ is sufficiently small, or your iteration budget is exceeded, then set $r^\star \leftarrow r_j$ and STOP.
  - If the sign of $g(r_j)$ matches that of $g(\ell_j)$, this means that $r^\star \in [r_j, u_j]$; so set $\ell_{j+1} \leftarrow r_j$ and $u_{j+1} \leftarrow u_j$. Otherwise, $r^\star \in [\ell_j, r_j]$; so set $\ell_{j+1} \leftarrow \ell_j$ and $u_{j+1} \leftarrow r_j$.

Each iteration of the algorithm chops the search area in two and therefore converges to $r^\star$ pretty quickly.

Georgia
Tech

Example: Use bisection to find $\sqrt{2}$ by solving $g(x) = x^2 - 2 = 0$.

In order to initialize the bisection algorithm, we note that $g(1) = -1$ and $g(2) = 2$. So there's a zero in $[1, 2]$ just itching to be found!

step	ℓ_j	$g(\ell_j)$	u_j	$g(u_j)$	r_j	$g(r_j)$
0	1	-1	2	2	1.5	0.25
1	1	-1	1.5	0.25	1.25	-0.4375
2	1.25	-0.4375	1.5	0.25	1.375	-0.1094
3	1.375	-0.1094	1.5	0.25	1.4375	0.0664
4	1.375	-0.1094	1.4375	0.0664	1.40625	-0.0225
$\vdots$						

You can see that r_j seems to be converging to $\sqrt{2} \doteq 1.4142$. $\qquad \square$

Newton's method. It's usually a lot faster than bisection. Here's a reasonable implementation.

1. Initialize r_0 as some first guess of the root. Set $j \leftarrow 0$.

2. Update $r_{j+1} \leftarrow r_j - g(r_j)/g'(r_j)$.

3. If $|g(r_{j+1})|$ or $|r_{j+1} - r_j|$ or your budget is suitably small, then STOP and set $r^\star \leftarrow r_{j+1}$. Otherwise, let $j \leftarrow j + 1$ and goto Step 2.

Example: Use Newton to find $\sqrt{2}$ by solving $g(x) = x^2 - 2 = 0$.

To do so, note that

$$r_{j+1} \leftarrow r_j - \frac{g(r_j)}{g'(r_j)} \;=\; r_{j+1} \leftarrow r_j - \frac{r_j^2 - 2}{2r_j} \;=\; \frac{r_j^2 + 2}{2r_j}.$$

If $r_0 = 1$, then we find that $r_1 = 3/2$, $r_2 = 17/12 \doteq 1.4167$, $r_3 = 577/408 \doteq 1.4142$, ... Wow, that's fast convergence! $\odot$

Georgia
Tech