18.440: Lecture 18

Uniform random variables

Scott Sheffield

MIT

Outline

Uniform random variable on $[0, 1]$

Uniform random variable on $[\alpha, \beta]$

Motivation and examples

Outline

Uniform random variable on $[0, 1]$

Uniform random variable on $[\alpha, \beta]$

Motivation and examples

Recall continuous random variable definitions

- Say X is a **continuous random variable** if there exists a **probability density function** $f = f_X$ on $\mathbb{R}$ such that $P\{X \in B\} = \int_B f(x)dx := \int 1_B(x)f(x)dx$.

Recall continuous random variable definitions

- Say X is a **continuous random variable** if there exists a **probability density function** $f = f_X$ on $\mathbb{R}$ such that $P\{X \in B\} = \int_B f(x)dx := \int 1_B(x)f(x)dx$.
- We may assume $\int_{\mathbb{R}} f(x)dx = \int_{-\infty}^{\infty} f(x)dx = 1$ and f is non-negative.

Recall continuous random variable definitions

- Say X is a **continuous random variable** if there exists a **probability density function** $f = f_X$ on $\mathbb{R}$ such that $P\{X \in B\} = \int_B f(x)dx := \int 1_B(x)f(x)dx$.
- We may assume $\int_{\mathbb{R}} f(x)dx = \int_{-\infty}^{\infty} f(x)dx = 1$ and f is non-negative.
- Probability of interval $[a, b]$ is given by $\int_a^b f(x)dx$, the area under f between a and b.

Recall continuous random variable definitions

- Say X is a **continuous random variable** if there exists a **probability density function** $f = f_X$ on $\mathbb{R}$ such that $P\{X \in B\} = \int_B f(x)dx := \int 1_B(x)f(x)dx$.
- We may assume $\int_{\mathbb{R}} f(x)dx = \int_{-\infty}^{\infty} f(x)dx = 1$ and f is non-negative.
- Probability of interval $[a, b]$ is given by $\int_a^b f(x)dx$, the area under f between a and b.
- Probability of any single point is zero.

Recall continuous random variable definitions

- Say X is a **continuous random variable** if there exists a **probability density function** $f = f_X$ on $\mathbb{R}$ such that $P\{X \in B\} = \int_B f(x)dx := \int 1_B(x)f(x)dx$.
- We may assume $\int_{\mathbb{R}} f(x)dx = \int_{-\infty}^{\infty} f(x)dx = 1$ and f is non-negative.
- Probability of interval $[a, b]$ is given by $\int_a^b f(x)dx$, the area under f between a and b.
- Probability of any single point is zero.
- Define **cumulative distribution function** $F(a) = F_X(a) := P\{X < a\} = P\{X \leq a\} = \int_{-\infty}^{a} f(x)dx$.

Uniform random variables on $[0,1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & x \notin [0,1]. \end{cases}$

Uniform random variables on $[0, 1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & x \notin [0,1]. \end{cases}$
- Then for any $0 \le a \le b \le 1$ we have $P\{X \in [a,b]\} = b - a$.

Uniform random variables on $[0, 1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & x \notin [0,1]. \end{cases}$
- Then for any $0 \le a \le b \le 1$ we have $P\{X \in [a, b]\} = b - a$.
- Intuition: all locations along the interval $[0, 1]$ equally likely.

Uniform random variables on $[0,1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & x \notin [0,1]. \end{cases}$
- Then for any $0 \le a \le b \le 1$ we have $P\{X \in [a,b]\} = b - a$.
- Intuition: all locations along the interval $[0,1]$ equally likely.
- Say that X is a **uniform random variable on** $[0,1]$ or that X is **sampled uniformly from** $[0,1]$.

Properties of uniform random variable on $[0, 1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & x \notin [0,1]. \end{cases}$

Properties of uniform random variable on $[0,1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & x \notin [0,1]. \end{cases}$
- What is $E[X]$?

Properties of uniform random variable on $[0,1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & x \notin [0,1]. \end{cases}$
- What is $E[X]$?
- Guess $1/2$ (since $1/2$ is, you know, in the middle).

Properties of uniform random variable on $[0,1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & x \notin [0,1]. \end{cases}$
- What is $E[X]$?
- Guess $1/2$ (since $1/2$ is, you know, in the middle).
- Indeed, $\int_{-\infty}^{\infty} f(x)xdx = \int_0^1 xdx = \frac{x^2}{2}\Big|_0^1 = 1/2$.

Properties of uniform random variable on $[0, 1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} 1 & x \in [0, 1] \\ 0 & x \notin [0, 1]. \end{cases}$
- What is $E[X]$?
- Guess $1/2$ (since $1/2$ is, you know, in the middle).
- Indeed, $\int_{-\infty}^{\infty} f(x) x dx = \int_0^1 x dx = \frac{x^2}{2}\Big|_0^1 = 1/2$.
- What would you guess the variance is? Expected square of distance from $1/2$?

Properties of uniform random variable on $[0, 1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & x \notin [0,1]. \end{cases}$
- What is $E[X]$?
- Guess $1/2$ (since $1/2$ is, you know, in the middle).
- Indeed, $\int_{-\infty}^{\infty} f(x)x dx = \int_0^1 x dx = \frac{x^2}{2}\Big|_0^1 = 1/2$.
- What would you guess the variance is? Expected square of distance from $1/2$?
- It's obviously less than $1/4$, but how much less?

Properties of uniform random variable on $[0,1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & x \notin [0,1]. \end{cases}$
- What is $E[X]$?
- Guess $1/2$ (since $1/2$ is, you know, in the middle).
- Indeed, $\int_{-\infty}^{\infty} f(x)xdx = \int_0^1 xdx = \frac{x^2}{2}\Big|_0^1 = 1/2$.
- What would you guess the variance is? Expected square of distance from $1/2$?
- It's obviously less than $1/4$, but how much less?
- $E[X^2] = \int_{-\infty}^{\infty} f(x)x^2dx = \int_0^1 x^2dx = \frac{x^3}{3}\Big|_0^1 = 1/3$.

Properties of uniform random variable on $[0,1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & x \notin [0,1]. \end{cases}$
- What is $E[X]$?
- Guess $1/2$ (since $1/2$ is, you know, in the middle).
- Indeed, $\int_{-\infty}^{\infty} f(x)x dx = \int_0^1 x dx = \frac{x^2}{2}\Big|_0^1 = 1/2$.
- What would you guess the variance is? Expected square of distance from $1/2$?
- It's obviously less than $1/4$, but how much less?
- $E[X^2] = \int_{-\infty}^{\infty} f(x)x^2 dx = \int_0^1 x^2 dx = \frac{x^3}{3}\Big|_0^1 = 1/3$.
- So $\mathrm{Var}[X] = E[X^2] - (E[X])^2 = 1/3 - 1/4 = 1/12$.

Properties of uniform random variable on $[0, 1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & x \notin [0,1]. \end{cases}$

Properties of uniform random variable on $[0,1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & x \notin [0,1]. \end{cases}$
- What is the cumulative distribution function $F_X(a) = P\{X < a\}$?

Properties of uniform random variable on $[0,1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & x \notin [0,1]. \end{cases}$
- What is the cumulative distribution function $F_X(a) = P\{X < a\}$?
- $F_X(a) = \begin{cases} 0 & a < 0 \\ a & a \in [0,1] \\ 1 & a > 1 \end{cases}.$

Properties of uniform random variable on $[0,1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & x \notin [0,1]. \end{cases}$
- What is the cumulative distribution function $F_X(a) = P\{X < a\}$?
- $F_X(a) = \begin{cases} 0 & a < 0 \\ a & a \in [0,1] \\ 1 & a > 1 \end{cases}.$
- What is the general moment $E[X^k]$ for $k \geq 0$?

Outline

Uniform random variable on $[0, 1]$

Uniform random variable on $[\alpha, \beta]$

Motivation and examples

Outline

Uniform random variable on $[0, 1]$

Uniform random variable on $[\alpha, \beta]$

Motivation and examples

Uniform random variables on $[\alpha, \beta]$

- Fix $\alpha < \beta$ and suppose X is a random variable with probability density function $f(x) = \begin{cases} \frac{1}{\beta-\alpha} & x \in [\alpha, \beta] \\ 0 & x \notin [\alpha, \beta]. \end{cases}$

Uniform random variables on $[\alpha, \beta]$

- Fix $\alpha < \beta$ and suppose X is a random variable with probability density function $f(x) = \begin{cases} \frac{1}{\beta-\alpha} & x \in [\alpha, \beta] \\ 0 & x \notin [\alpha, \beta]. \end{cases}$
- Then for any $\alpha \leq a \leq b \leq \beta$ we have $P\{X \in [a, b]\} = \frac{b-a}{\beta-\alpha}$.

Uniform random variables on $[\alpha, \beta]$

- Fix $\alpha < \beta$ and suppose X is a random variable with probability density function $f(x) = \begin{cases} \frac{1}{\beta-\alpha} & x \in [\alpha, \beta] \\ 0 & x \notin [\alpha, \beta]. \end{cases}$
- Then for any $\alpha \le a \le b \le \beta$ we have $P\{X \in [a, b]\} = \frac{b-a}{\beta-\alpha}$.
- Intuition: all locations along the interval $[\alpha, \beta]$ are equally likely.

Uniform random variables on $[\alpha, \beta]$

- Fix $\alpha < \beta$ and suppose X is a random variable with probability density function $f(x) = \begin{cases} \frac{1}{\beta-\alpha} & x \in [\alpha,\beta] \\ 0 & x \notin [\alpha,\beta]. \end{cases}$
- Then for any $\alpha \le a \le b \le \beta$ we have $P\{X \in [a,b]\} = \frac{b-a}{\beta-\alpha}$.
- Intuition: all locations along the interval $[\alpha,\beta]$ are equally likely.
- Say that X is a **uniform random variable on** $[\alpha,\beta]$ or that X is **sampled uniformly from** $[\alpha,\beta]$.

Properties of uniform random variable on $[0, 1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} \frac{1}{\beta - \alpha} & x \in [\alpha, \beta] \\ 0 & x \notin [\alpha, \beta]. \end{cases}$

Properties of uniform random variable on $[0,1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} \frac{1}{\beta-\alpha} & x \in [\alpha,\beta] \\ 0 & x \notin [\alpha,\beta]. \end{cases}$
- What is $E[X]$?

Properties of uniform random variable on $[0,1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} \frac{1}{\beta-\alpha} & x \in [\alpha,\beta] \\ 0 & x \notin [\alpha,\beta]. \end{cases}$
- What is $E[X]$?
- Intuitively, we'd guess the midpoint $\frac{\alpha+\beta}{2}$.

Properties of uniform random variable on $[0,1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} \frac{1}{\beta-\alpha} & x \in [\alpha,\beta] \\ 0 & x \notin [\alpha,\beta]. \end{cases}$
- What is $E[X]$?
- Intuitively, we'd guess the midpoint $\frac{\alpha+\beta}{2}$.
- What's the cleanest way to prove this?

Properties of uniform random variable on $[0,1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} \frac{1}{\beta-\alpha} & x \in [\alpha,\beta] \\ 0 & x \notin [\alpha,\beta]. \end{cases}$
- What is $E[X]$?
- Intuitively, we'd guess the midpoint $\frac{\alpha+\beta}{2}$.
- What's the cleanest way to prove this?
- One approach: let Y be uniform on $[0,1]$ and try to show that $X = (\beta-\alpha)Y + \alpha$ is uniform on $[\alpha,\beta]$.

Properties of uniform random variable on $[0,1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} \frac{1}{\beta-\alpha} & x \in [\alpha,\beta] \\ 0 & x \notin [\alpha,\beta]. \end{cases}$
- What is $E[X]$?
- Intuitively, we'd guess the midpoint $\frac{\alpha+\beta}{2}$.
- What's the cleanest way to prove this?
- One approach: let Y be uniform on $[0,1]$ and try to show that $X = (\beta-\alpha)Y + \alpha$ is uniform on $[\alpha,\beta]$.
- Then linearity of $E[X] = (\beta-\alpha)E[Y] + \alpha = (1/2)(\beta-\alpha) + \alpha = \frac{\alpha+\beta}{2}$.

Properties of uniform random variable on $[0, 1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} \frac{1}{\beta-\alpha} & x \in [\alpha, \beta] \\ 0 & x \notin [\alpha, \beta]. \end{cases}$
- What is $E[X]$?
- Intuitively, we'd guess the midpoint $\frac{\alpha+\beta}{2}$.
- What's the cleanest way to prove this?
- One approach: let Y be uniform on $[0, 1]$ and try to show that $X = (\beta - \alpha)Y + \alpha$ is uniform on $[\alpha, \beta]$.
- Then linearity of $E[X] = (\beta - \alpha)E[Y] + \alpha = (1/2)(\beta - \alpha) + \alpha = \frac{\alpha+\beta}{2}$.
- Using similar logic, what is the variance $\mathrm{Var}[X]$?

Properties of uniform random variable on $[0,1]$

- Suppose X is a random variable with probability density function $f(x) = \begin{cases} \frac{1}{\beta-\alpha} & x \in [\alpha,\beta] \\ 0 & x \notin [\alpha,\beta]. \end{cases}$
- What is $E[X]$?
- Intuitively, we'd guess the midpoint $\frac{\alpha+\beta}{2}$.
- What's the cleanest way to prove this?
- One approach: let Y be uniform on $[0,1]$ and try to show that $X = (\beta-\alpha)Y + \alpha$ is uniform on $[\alpha,\beta]$.
- Then linearity of $E[X] = (\beta-\alpha)E[Y] + \alpha = (1/2)(\beta-\alpha) + \alpha = \frac{\alpha+\beta}{2}$.
- Using similar logic, what is the variance $\mathrm{Var}[X]$?
- Answer: $\mathrm{Var}[X] = \mathrm{Var}[(\beta-\alpha)Y + \alpha] = \mathrm{Var}[(\beta-\alpha)Y] = (\beta-\alpha)^2\mathrm{Var}[Y] = (\beta-\alpha)^2/12$.

Outline

Uniform random variable on $[0, 1]$

Uniform random variable on $[\alpha, \beta]$

Motivation and examples

Outline

Uniform random variable on $[0, 1]$

Uniform random variable on $[\alpha, \beta]$

Motivation and examples

Uniform random variables and percentiles

- Toss $n = 300$ million Americans into a hat and pull one out uniformly at random.

Uniform random variables and percentiles

- Toss $n = 300$ million Americans into a hat and pull one out uniformly at random.
- Is the height of the person you choose a uniform random variable?

Uniform random variables and percentiles

- Toss $n = 300$ million Americans into a hat and pull one out uniformly at random.
- Is the height of the person you choose a uniform random variable?
- Maybe in an approximate sense?

Uniform random variables and percentiles

- Toss $n = 300$ million Americans into a hat and pull one out uniformly at random.
- Is the height of the person you choose a uniform random variable?
- Maybe in an approximate sense?
- No.

Uniform random variables and percentiles

- Toss $n = 300$ million Americans into a hat and pull one out uniformly at random.
- Is the height of the person you choose a uniform random variable?
- Maybe in an approximate sense?
- No.
- Is the *percentile* of the person I choose uniformly random? In other words, let p be the fraction of people left in the hat whose heights are less than that of the person I choose. Is p, in some approximate sense, a uniform random variable on $[0, 1]$?

Uniform random variables and percentiles

- Toss $n = 300$ million Americans into a hat and pull one out uniformly at random.
- Is the height of the person you choose a uniform random variable?
- Maybe in an approximate sense?
- No.
- Is the *percentile* of the person I choose uniformly random? In other words, let p be the fraction of people left in the hat whose heights are less than that of the person I choose. Is p, in some approximate sense, a uniform random variable on $[0, 1]$?
- The way I defined it, p is uniform from the set $\{0, 1/(n-1), 2/(n-1), \ldots, (n-2)/(n-1), 1\}$. When n is large, this is kind of like a uniform random variable on $[0, 1]$.

Approximately uniform random variables

- Intuition: which of the following should give approximately uniform random variables?

Approximately uniform random variables

- Intuition: which of the following should give approximately uniform random variables?
- 1. Toss $n = 300$ million Americans into a hat, pull one out uniformly at random, and consider that person's height (in centimeters) modulo one.

Approximately uniform random variables

- Intuition: which of the following should give approximately uniform random variables?
- 1. Toss $n = 300$ million Americans into a hat, pull one out uniformly at random, and consider that person's height (in centimeters) modulo one.
- 2. The location of the first raindrop to land on a telephone wire stretched taut between two poles.
- 3. The amount of time you have to wait until the next subway train come (assuming trains come promptly every six minutes and you show up at kind of a random time).

Approximately uniform random variables

- Intuition: which of the following should give approximately uniform random variables?
- 1. Toss $n = 300$ million Americans into a hat, pull one out uniformly at random, and consider that person's height (in centimeters) modulo one.
- 2. The location of the first raindrop to land on a telephone wire stretched taut between two poles.
- 3. The amount of time you have to wait until the next subway train come (assuming trains come promptly every six minutes and you show up at kind of a random time).
- 4. The amount of time you have to wait until the next subway train (without the parenthetical assumption above).

Approximately uniform random variables

- 5. How about the location of the jump between times 0 and 1 of λ-Poisson point process (which we condition to have exactly one jump between $[0, 1]$)?

Approximately uniform random variables

- 5. How about the location of the jump between times 0 and 1 of λ-Poisson point process (which we condition to have exactly one jump between $[0, 1]$)?
- 6. The location of the ace of spades within a shuffled deck of 52 cards.