Lecture 22

Uniform distributions

Probability densities

Infinitesimals

Expectation

Discrete $\longrightarrow$ continuous

Expectation as prediction

Uniform distributions

- Tossing a die: $P\{\text{Number} = k\} = 1/6\}$.
- Drawing a number at random from $\{1, 2, \ldots, n\}$:

$$\begin{aligned} P\{\text{Number} = k\} &= \frac{1}{n} \\ P\{\text{Number} \in A\} &= \frac{\#(A)}{n} \end{aligned}$$

- Drawing a number at random from $[0, 1]$:

$$P\{\text{Number} \in A\} = |A|.$$

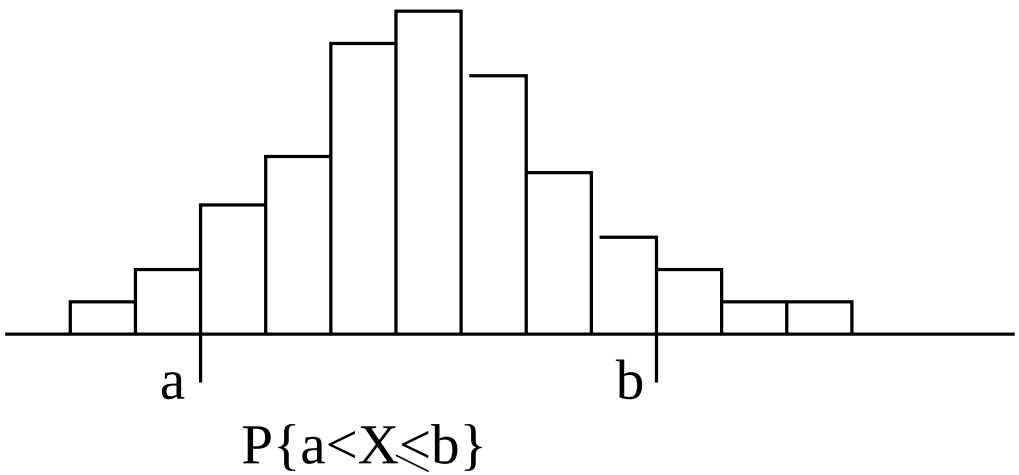
- Drawing a number at random from $[0, 2]$:

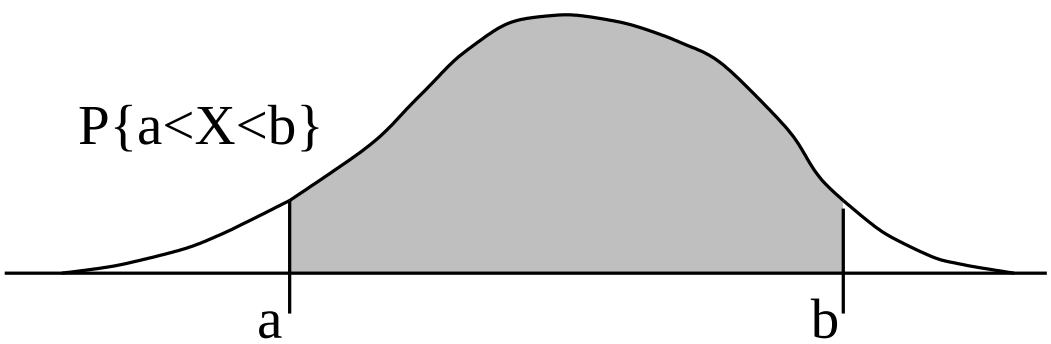
$$P\{\text{Number} \in A\} = |A|/2.$$

- Drawing a point at random from $S =$ Unit disk:

$$P\{\text{Point} \in A\} = |A \cap S|/\pi.$$

From discrete to continuous





Densities

- A real valued random variable X has density f $\iff$

$$
\begin{aligned}
P\{a \le X \le b\} &= \int_a^b f(x)\,dx, \\
P\{X \in A\} &= \int_A f(x)\,dx,
\end{aligned}
$$

for all numbers $a \le b$ and set $A \subset \Re$.

- Immediate properties:

  - $f(x) \ge 0, \quad$ for all x.
  - $\int_{-\infty}^{\infty} f(x)\,dx \ = 1.$
  - Do NOT need $f(x) \le 1$!
  - $P\{X = x\} \ = \ 0$ for all x.

Infinitesimals

$$
\begin{aligned}
P\{a \le X \le b\} &= \int_a^b f(x)\,dx, \\
\Rightarrow \qquad P\{x \le X \le x + dx\} &= \int_x^{x+dx} f(x)\,dx \;\approx\; f(x)\,dx \\
\text{and} \qquad P\{x - dx \le X \le x\} &= \int_{x-dx}^{x} f(x)\,dx \;\approx\; f(x)\,dx
\end{aligned}
$$

- We write this as

$$\boxed{P\{X \in dx\} \;=\; f(x)\,dx}$$

FROM DISCRETE TO CONTINUOUS

Concept	Discrete	Continuous
Probability	$P\{X = x\} = p_X(x)$	$P\{X \in dx\} = f_X(x)\,dx$
Mode	argmax $p_X(x)$	argmax$f_X(x)$
Expectation	$\sum_x x\,p_X(x)$	$\int_{-\infty}^{\infty} x\,f_X(x)\,dx$
$Y = g(X)$	$p_Y(y) = \sum_{x:g(x)=y} p_X(x)$	???
$E\{g(X)\}$	$\sum_x g(x)\,p_X(x)$	$\int_{-\infty}^{\infty} g(x)\,f_X(x)\,dx$
Mean	$\sum_x x\,p_X(x)$	$\int_{-\infty}^{\infty} x f_X(x)\,dx$
Variance	$E\{[X - \mu_X]^2\}$	$E\{[X - \mu_X]^2\}$
Indepen-dence	$P\{X \in A, Y \in B\} = P\{X \in A\}P\{Y \in B\}$	$P\{X \in A, Y \in B\} = P\{X \in A\}P\{Y \in B\}$
“	$p_{XY}(x, y) = p_X(x)p_Y(y)$	???

Justifying $E\{X\} = \int_{-\infty}^{\infty} x\, f_X(x)\, dx$

- Take X continuous with density f and values in [0,1].
- $\iff \quad f(x) = 0$ if $x \notin [0,1]$.
- Fix $n \geq 1$ and define a discrete approximation X^n to X:

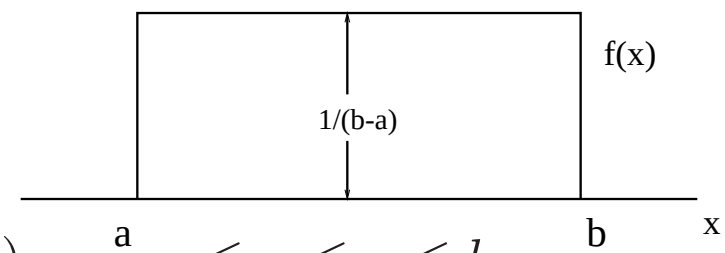
$$X \in \left[\frac{k}{n}, \frac{k+1}{n}\right) \quad \Rightarrow X^n = \frac{k}{n}$$

- Then $|X^n - X| \leq 1/n$
- $P\{X^n = k/n\} = P\{k/n \leq X < (k+1)/n\} \approx f(k/n)/n$

$$\begin{aligned} E\{X^n\} &= \sum_{k=0}^{n-1} (k/n) P\{X^n = k/n\} \approx \sum_{k=0}^{n-1} (k/n) f(k/n)(1/n) \\ &\approx \int_0^1 x\, f(x)\, dx \equiv E\{X\} \end{aligned}$$

1: $X \sim U[a,b]$: Uniform distribution on $[a,b]$.

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \le x \le b \ , \\ 0 & \text{otherwise.} \end{cases}$$



- $P\{x \le X \le y\} = (y-x)/(b-a), \quad a \le x \le y \le b.$
- Standardisation:

$$X \sim U[a,b] \iff U \equiv \frac{X-a}{b-a} \sim U[0,1]$$

PROOF: Check the probabilities, (for $0 \le x < y \le 1$).

$$\begin{aligned} P\{x \le U \le y\} &= P\{x \le (X-a)/(b-a) \le y\} \\ &= P\{a+(b-a)x \le X \le a+(b-a)y\} \\ &= \frac{(b-a)(y-x)}{b-a} = y-x \end{aligned}$$

Mean of $X \sim U[a,b]$

- Directly:

$$E\{X\} \;=\; \int_a^b x\,\frac{1}{b-a}\,dx \;=\; \frac{b^2-a^2}{2(b-a)} \;=\; \frac{b+a}{2}$$

- Via standardisation:

- $U \sim U[0,1] \iff X = a + (b-a)U \sim U[a,b]$.
- $\Rightarrow\; E\{X\} \;=\; E\{a+(b-a)U\} = a + (b-a)E\{U\}$
- $E\{U\} \;= \int_0^1 u\,du \;=\; [u^2/2]_0^1 \;=\; 1/2$
- $\Rightarrow\; E\{X\} \;=\; a + (b-a)/2 \;=\; (b+a)/2$

Variance of $X \sim U[a,b]$

- Recall: $Var(aX+b) = a^2 Var(X)$.
- $X = a + (b-a)U \;\Rightarrow\; Var(X) = (b-a)^2 Var(U)$
- $Var(U) = E\{U^2\} - (E\{U\})^2 \;=\; E\{U^2\} - 1/4$

$$\begin{aligned} E\{U^2\} &= \int_{-\infty}^{\infty} u^2 \, f_U(u) \, du \\ &= \int_0^1 u^2 \, du \\ &= \left[\frac{u^3}{3}\right]_0^1 \;=\; 1/3 \end{aligned}$$

- $Var(U) \;=\; 1/3 - 1/4 \;=\; 1/12$
- $SD(X) \;=\; (b-a)/\sqrt{12}$

Rules from the discrete case

- $E\{aX + b\} \;=\; aE\{X\} + b$
- $Var(aX + b) \;=\; a^2 Var(X)$
- $\Rightarrow SD(aX + b) \;=\; aSD(X)$
- and also, if

$$X^* = \frac{X - E\{X\}}{SD(X)}$$

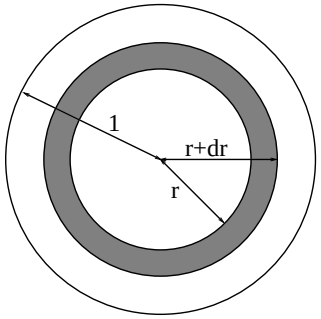
then

$$E\{X^*\} = 0, \quad \text{and} \quad Var(X^*) = 1$$

- $E\{X_1 + \ldots + X_n\} \;=\; E\{X_1\} + \ldots + E\{X_n\}$
- Similar things for $E\{g(X)\}$, $E\{XY\}$, $Var(X + Y)$, etc.

BECAUSE EXPECTATION IS A LINEAR OPERATION

Radial distance



Choose a point at random from the unit disk.
R is its distance from the origin.

$$\begin{aligned} P\{R \in dr\} &= P\{r \le R \le r + dr\} \\ &= \frac{\text{Area of annulus}}{\text{Total area}} = \frac{\pi(r+dr)^2 - \pi r^2}{\pi} \sim 2r\ dr \end{aligned}$$

$$\Rightarrow \qquad f(r) = 2r \qquad (\text{if } 0 < r < 1)$$

$$\begin{aligned} E\{R\} &= \int_{-\infty}^{\infty} r\ f(r)\ dr = \int_0^1 2r^2\ dr = \frac{2}{3}\Big[r^3\Big] - 0^1 = \frac{2}{3} \\ E\{R^2\} &= \int_{-\infty}^{\infty} r^2\ f(r)\ dr = \int_0^1 2r^3\ dr = \frac{2}{4}\Big[r^4\Big] - 0^1 = \frac{1}{2} \\ Var(R) &= E\{R^2\} - (E\{R\})^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18} \end{aligned}$$

Expectation and Prediction (Lec 16)

- AIM: To predict a value of X.
- LOSS FUNCTION: $L(x, b)$ measures the "cost" of the guess b if the real value of X is x.
- POLICY: Minimise $E\{L(X, b)\}$.

Minimising absolute error

$$L(x, b) \;=\; |x - b|$$

Note: $$\frac{d}{db}|x-b| = \begin{cases} 1 & \text{if } x < b, \\ -1 & \text{if } x > b. \end{cases}$$

$$\begin{aligned}
\Rightarrow \qquad \frac{d}{db}E\{L(X,b)\} &= \frac{d}{db}\int_{-\infty}^{\infty} |x-b| f(x)\,dx \\
&= \int_{x \le b} f(x)\,dx \;-\; \int_{x > b} f(x)\,dx \\
&= P\{X \le b\} \;-\; P\{X > b\} \\
&= 2P\{X \le b\} \;-\; 1 \\
&= 0, \qquad \text{if } b \text{ is the } \underline{\text{median}} \text{ of } X.
\end{aligned}$$

$\Rightarrow$ Under "absolute error" a median is the best guess.