10/15

TODAY

1. Review of Expected values

2. VARIANCE / Standard Dev.

3. Entropy

   ↳ measure randomness

FRIDAY

Review

Next week

Midterm

Easy Expected Values

Looking at a fair/Balanced die ⇝ possible outcomes

$\Omega = \{1, 2, 3, 4, 5, 6\}$

Calculate expected value

$$\sum_{i \in \Omega} v(i)\, p(i) = \frac{1}{6}\cdot 1 + \frac{1}{6}\cdot 2 + \frac{1}{6}\cdot 3 + \frac{1}{6}\, 4 + \frac{1}{6}\cdot 5 + \frac{1}{6}\cdot 6$$

$$= 3.5$$

What is the expected "Sum" when we roll a pair of Dice?

Die 1 \ Die 2	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

2, 3, 3, 4, 4, 4, 5, 5, 5, 5,
6, 6, 6, 6 . . . . . 11, 11, 12

$$\sum_{i \in R} v(i)\, p(i) \qquad + 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + 5\, \frac{4}{36} + 6\, \frac{5}{36} + 7\, \frac{6}{36} + 8\, \frac{5}{36} + 9\, \frac{4}{36} + 10 \cdot \frac{5}{36} + 11\, \frac{6}{36} + 12 \cdot \frac{?}{36}$$

$$= \frac{252}{36}$$

$$= 7$$

The roll of Die are INDEPENDANT (Both fair) so $E(x_1 + x_2) = E(x_1) + E(x_2)$

Expected value considers distribution of outcomes.

Rolling a pair of dice

→ more 7's than 3's

---

what are the "properties" of the list (distribution)

Ex Test Scores

Exam 1 Set 1 75 75 76 77 78 79 80 80 81 82 83 84 85 85

Exam 2 Set 2 60 61 62 62 62 63 63 97 97 98 98 98 99 100

We can measure how far away a given value, x, is from the expected value, $E[x]$.

Ex 2

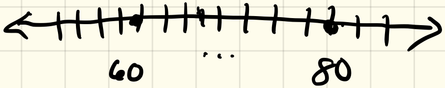
Exam 1	Exam 2	Exam 3
60, 80, 100	79, 80, 81	75, 80, 85
$E[x]=80$	$E[x]=80$	$E[x]=80$
$60-80=-20$	$79-80=-1$	$75-80=-5$
$80-80=0$	$80-80=0$	$80-80=0$
$100-80=20$	$81-80=1$	$85-80=5$
$\frac{-20+0+20}{3}=\frac{0}{3}=0$	$\frac{-1+0+1}{3}=0$	$\frac{-5+0+5}{3}=0$

Just take the difference

It doesn't work.

Find the "distance" to expected value.

60	79	75
80	80	80
100	81	85



Distance from expected value.

called Mean absolute deviation

$|60-80| = 20$

$|80-80| = 0$

$|100-80| = 20$

$\frac{40}{3}$

$|79-80| = 1$

$|80-80| = 0$

$|81-80| = 1$

$\frac{2}{3}$

$|75-80| = 5$

$|80-80| \quad 0$

$|85-80| \quad 5$

Mean absolute deviation $= \sum_{x \in \Omega} \frac{|x - E[x]|}{|\Omega|}$

sample size $\frac{10}{3}$

Square the differences — Amplification

Big gets bigger

Small (<1) get smaller

Variance

$$\sigma^2 = \frac{\sum_{x \in \Omega} (x - E[x])^2}{|\Omega|}$$

sigma

So for our example

$E[x^2]$

20^2	1^2	5^2
0^2	0^2	0^2
20^2	1^2	5^2
$800/3$	$2/3$	$50/3$

~ wrong units

STANDARD DEVIATION

We can take square root

$$\sigma = \sqrt{\sum_{x \in \Omega} \frac{(x - E(x))^2}{|\Omega|}}$$

Measure Randomness

Consider an example of encoding letters

ASCII
Fixed
length encoding
1 character (letter)
= 8 bits

Smaller example

$\{A, B, C, D, E, F, G, H\}$

↳ we could do a fixed 3 bit encoding

A − 000
B − 001

C = 010
D = 011

E = 100
F = 101

G = 110
H = 111

~ All numbers have 3 bits −
But what if one letter was really common?

What if we allowed variable lengths for letters

→ Common letters - short codes

→ rare letters — long codes

Could we be more efficient?

More formally, BPC - Bits per code

What if we let the length depend on frequency

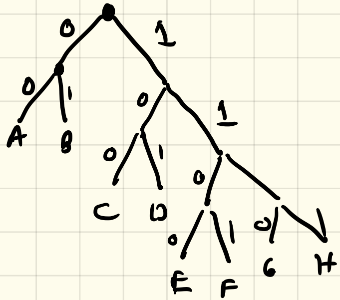
$$BPC = \sum_x r(x)\,p(x)$$

$$= \sum_i \ell(i)\,p(i)$$

$\ell(i)$ - length of i^{th} character

$p(i)$ - probability of i^{th} character

letter	A	B	C	D	E	F	G	H
probability	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
Consider this encoding	00	01	100	101	1100	1101	1110	1111



But is this better

$$BPC = \sum_i l(i)\,p(i)$$

$$2\cdot\frac{1}{4} + 2\frac{1}{4} + 3\frac{1}{8} + 3\frac{1}{8} + 4\frac{1}{16} + 4\cdot\frac{1}{16} + 4\frac{1}{16} + 4\frac{1}{16} \cong 2.75$$

Picked on purpose

$$p_i = \left(\frac{1}{2}\right)^k$$

$$l_i = k$$

observation from our example

$$p_i = \left(\frac{1}{2}\right)^{l_i}$$

substitution

$$2^{l_i} = \frac{1}{p_i}$$

$$l_i = \lg\left(\frac{1}{p_i}\right)$$

$\lg$ — $\log_2$

Entropy — compression limit with this encoding

A measure of randomness

Expected code length

$$\sum_i p_i \lg\left(\frac{1}{p_i}\right) =$$

$$-\sum_i p_i \lg(p_i)$$ — entropy

$$\lg\left(\frac{1}{n}\right) = -\lg(n)$$

---

f - be a function that outputs the description of something

f can have input

$$d = f(x)$$

$$d^* = f(x)$$

$$K(d) = \min\left\{ \|f(x)\| \;\; \begin{array}{l} f \text{ is function} \\ \text{s.t } f(x) \text{ outputs } d \end{array} \right\}$$

Kolmogorov complexity — not solvable