# Variance

## Consider heights

| case 1 | case 2 | case 3 |
| --- | --- | --- |
| 4'10'' 5' 5'2'' | 4' 5' 6' | 3' 5' 7' |
| $E\{X_1\} = 5'$ | $E\{X_2\} = 5'$ | $E\{X_3\} = 5'$ |

How to measure "variability"

3 ways

① $Y = X - E\{X\}$ ✗ neg. & pos. cancel

② $Y = |X - E\{X\}|$ - mean absolute deviation

③ $Y = (X - E\{X\})^2$ - variance

case 2

① $E[Y] = \sum_{\omega \in \Omega} Y(\omega) \cdot p(\omega)$

$= (-12'') \cdot 1/3 + (0'') \cdot 1/3 + (+12'') \cdot 1/3$

$= 0$

② $E\{Y\} = |-12''| \cdot 1/3 + |0''| \cdot 1/3 + |+12''| \cdot 1/3$

$= 12 \cdot 1/3 + 0 \cdot 1/3 + 12 \cdot 1/3 = 8$ inches

case 1

③ $E\{Y\} = E[(X - E\{X\})^2] = (-2'')^2 \cdot 1/3 + (0'')^2 \cdot 1/3 + (+2'')^2 \cdot 1/3$

$= 4/3 + 0 + 4/3 = 8/3 \text{ inches}^2$ variance $\sigma^2$

Take square root, get back inches...

$$\sigma^2 = Var(X) = E\{(X - E\{X\})^2\}$$

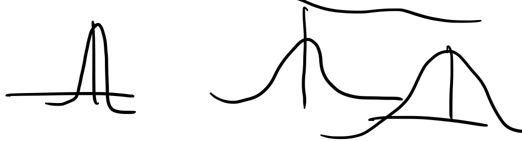
$$\Rightarrow \quad \sigma = \sqrt{Var(X)} = \sqrt{E[(X - E\{X\})^2]}$$ – back in units originally measured.

e.g.

case 1 : $\sigma^2 = Var(X) = 8/3 \text{ inches}^2$

$$\sigma = \sqrt{Var(X)} = \sqrt{8/3 \text{ inches}^2} = 1.63''$$

case 3 : ... $\sigma = 19.6''$



Exam

$E\{X\} = 76$

$\sigma = 15$ $\qquad \pm 2/3\,\sigma$ – quartiles

$2/3\,\sigma = 10$ pts

# Entropy

- Consider 8 letters only $\{A, B, C, D, E, F, G, H\}$
- Need 3-bits

$$\left.\begin{array}{l} A \to 000 \\ B \to 001 \\ \vdots \\ H \to 111 \end{array}\right\} \text{fixed length code}$$

- suppose must encode $n$-things, need: $\lceil \log_2 n \rceil$
- Efficiency of code is measured in bits-per-character on average, BPC. $BPC = 3$
- variable length code: assign short codes to frequent letters
  long codes to infrequent letters.
- $$BPC = \sum_{w \in \Omega} X(w) \cdot p(w) = \sum_i l_i \cdot p_i$$

$l_i$: code length for $i^{th}$ letter

$p_i$: probability of $i^{th}$ letter

Why must we have long codes to make up for short codes?

Analog to PHP: Kraft's Inequality

Example:

| A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|
| 00 | 01 | 010 | 011 | 100 | 101 | 110 | 111 |

BBB → 010101

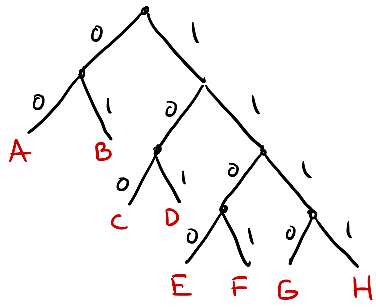
CF → 010101

Suppose we have following skewed distribution

$$P_1 = \begin{array}{cccccccc} A & B & C & D & E & F & G & H \\ (1/4, & 1/4, & 1/8, & 1/8, & 1/16, & 1/16, & 1/16, & 1/16) \end{array}$$

$$(00,\ 01,\ 100,\ 101,\ 1100,\ 1101,\ 1110,\ 1111)$$

Code Tree:




ADF → 001011101

$$\underset{A}{00}\,|\,\underset{D}{101}\,|\,\underset{F}{1101}$$

$$BPC = \sum_i l_i \cdot p_i$$

$$= 2 \cdot 1/4 + 2 \cdot 1/4 + 3 \cdot 1/8 + 3 \cdot 1/8 + 4 \cdot 1/16 + \dots + 4 \cdot 1/16$$

$$= 2.75$$

* 8.33% savings over fixed code of length 3.

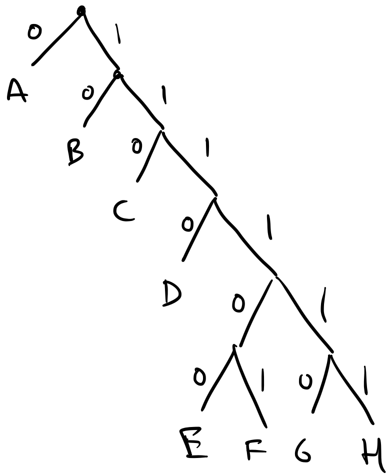
| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| $p_2 =$ | $\frac{1}{2}$ | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{16}$ | $\frac{1}{64}$ | $\frac{1}{64}$ | $\frac{1}{64}$ | $\frac{1}{64}$ |
| | 0 | 10 | 110 | 1110 | 111100 | 111101 | 111110 | 111111 |

$$BPC = \sum_i l_i \cdot p_i$$

$$= 1 \cdot \tfrac{1}{2} + 2 \cdot \tfrac{1}{4} + 3 \cdot \tfrac{1}{8} + 4 \cdot \tfrac{1}{16}$$

$$+ 6 \cdot \tfrac{1}{64} \cdots + 6 \cdot \tfrac{1}{64}$$

$$= 2$$




| $p_i$ | $l_i$ |
|---|---|
| $\frac{1}{2}$ | 1 |
| $\frac{1}{4}$ | 2 |
| $\frac{1}{8}$ | 3 |
| $\frac{1}{16}$ | 4 |
| $\frac{1}{64}$ | 6 |

$$p_i = \frac{1}{2^k}, \quad l_i = k \implies p_i = \frac{1}{2^{l_i}}$$

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

$$\iff \boxed{l_i = \log_2 \frac{1}{p_i}}$$

# Entropy.—

Q: $P = (p_1, p_2, p_3, \ldots)$

$$l_i = \log_2 \frac{1}{p_i}$$

What is expected code length?

$$BPC = \sum_i p_i \cdot l_i$$

$$= \boxed{\sum_i p_i \cdot \log_2 \frac{1}{p_i}}$$ entropy

$$H(p) = \sum_i p_i \cdot \log_2 \frac{1}{p_i}$$

ACBADAFGAH

$P_A = \frac{4}{10}$

$P_B = 1/10$

$P_C = 1/10$

$P_D = 1/10$

$\vdots$

**Facts**

1. Entropy is the compression limit.
2. Entropy is essential to bound an ability to communicate in presence of noise.
3. Fundamental measure of randomness.
4. Cryptography