Lecture 17

Variance

Standard deviation

Properties

Chebychev's inequality

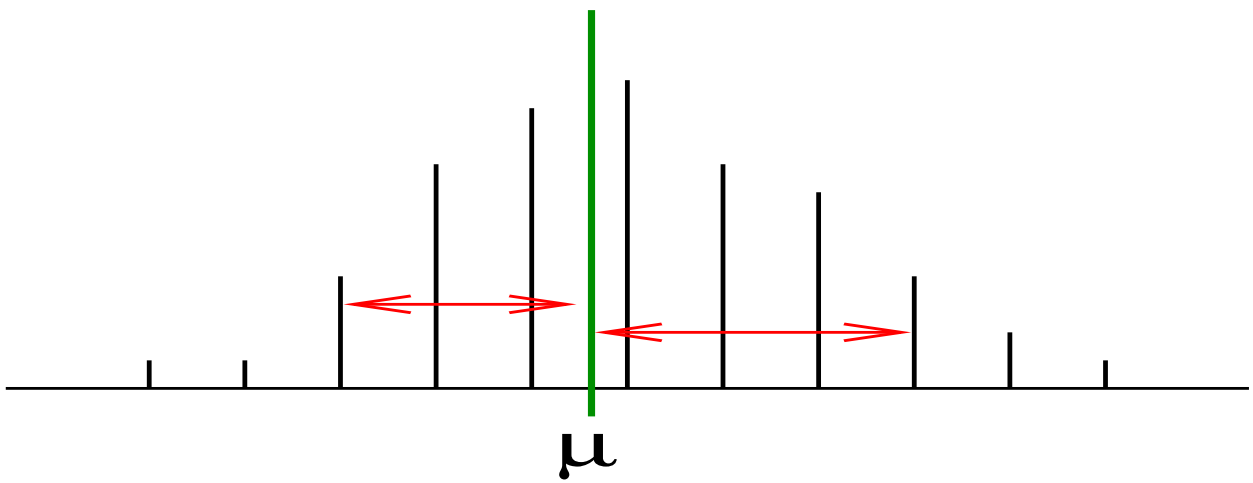
Variance:

$$Var(X) \equiv E\{(X-\mu)^2\}$$

$\mu = E\{X\}$

Standard deviation:

$$SD(X) \equiv \sqrt{Var(X)}$$



Computational formula:

$$Var(X) \;=\; E\{X^2\} \;-\; [E\{X\}]^2$$

$$\begin{aligned}
Var(X) &= E\{(X-\mu)^2\} \\
&= E\{X^2 \;-\; 2\mu X \;+\; \mu^2\} \\
&= E\{X^2\} \;-\; E\{2\mu X\} \;+\; E\{\mu^2\} \\
&= E\{X^2\} \;-\; 2\mu E\{X\} \;+\; \mu^2 \\
&= E\{X^2\} \;-\; \mu^2
\end{aligned}$$

1: Sampling from $\{1, 2, \ldots, n\}$

- X = number sampled
- $\mu = E\{X\} = \frac{1}{n}\sum_{k=1}^{n} k = (n+1)/2$

$$
\begin{aligned}
E\{X^2\} = \frac{1}{n}\sum_{k=1}^{n} k^2 &= \frac{1}{n}\frac{n(n+1)(2n+1)}{6} \\
&= \frac{(n+1)(2n+1)}{6} \\
\Rightarrow \qquad Var(X) &= \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 \\
&= \frac{n^2-1}{12} \\
\Rightarrow \qquad SD(X) &= \sqrt{\frac{n^2-1}{12}} \approx \frac{n}{3.46} \quad \text{for large } n.
\end{aligned}
$$

2: Indicators

- $I_A = 1$ if A occurs, and 0 if it does not.
- $E\{I_A\} \;=\; 0 \times P\{A^c\} + 1 \times P\{A\} \;=\; P\{A\} \;\equiv\; p.$
- $E\{I_A^2\} \;=\; 0^2 \times P\{A^c\} + 1^2 \times P\{A\} \;=\; P\{A\} \;=\; p.$
- $Var(I_A) \;=\; p - p^2 \;=\; p(1-p) \;=\; pq.$
- $SD(I_A) \;=\; \sqrt{pq}.$

3: Binomial variable

- $Var(X) \;= \sum_{k=0}^{n} k^2 \binom{n}{k} p^k q^{n-k} \;-\; (np)^2$
  $. \qquad\qquad = \; npq + n^2p^2 - (np)^2 \;=\; npq.$

Properties of $Var(X)$ and $SD(X)$.

1: Scaling and shifting

- $Var(aX + b) \;=\; a^2\, Var(X)$.
- $SD(aX + b) \;=\; a\, SD(X)$.

Set $\mu = E\{X\}$.

$$\begin{aligned} Var(aX+b) &= E\{[(aX+b) - E\{(aX+b)\}]^2\} \\ &= E\{[aX + b - (a\mu + b)]^2\} \\ &= E\{a^2[X-\mu]^2\} \\ &= a^2 Var(X) \end{aligned}$$

- Example: Celcius (X) to Fahrenheit (Y) via $Y = 32 + \frac{9}{5}X$ implies that $SD(Y) = \frac{9}{5}SD(X)$, and the "32" has no effect at all.

2: Standardisation

- Suppose X has mean μ and standard deviation $\sigma > 0$. Then

$$X^* \equiv \frac{X-\mu}{\sigma}$$

  is called the <u>standardised value</u> of X, or X in <u>standard units</u>.

- $E\{X^*\} = 0, \qquad Var(X^*) = 1.$

PROOF:

$$\begin{aligned} E\{aX+b\} &= a\mu + b, \quad \text{and} \quad Var(aX+b) = a^2\sigma^2 \\ \Rightarrow \qquad E\{X^*\} &= E\left\{\frac{1}{\sigma}X + \frac{-\mu}{\sigma}\right\} = 0 \\ \text{and} \qquad Var(X^*) &= Var\left(\frac{1}{\sigma}X + \frac{-\mu}{\sigma}\right) = \frac{1}{\sigma^2}\,\sigma^2 = 1 \end{aligned}$$

2.1: Uses of standardisation

- Comparison of <u>like</u> measurements.

  - e.g. Standardising Celcius and Fahrenheit measurements taken on a random day puts them on the same scale.

- Comparison of <u>unlike</u> measurements.

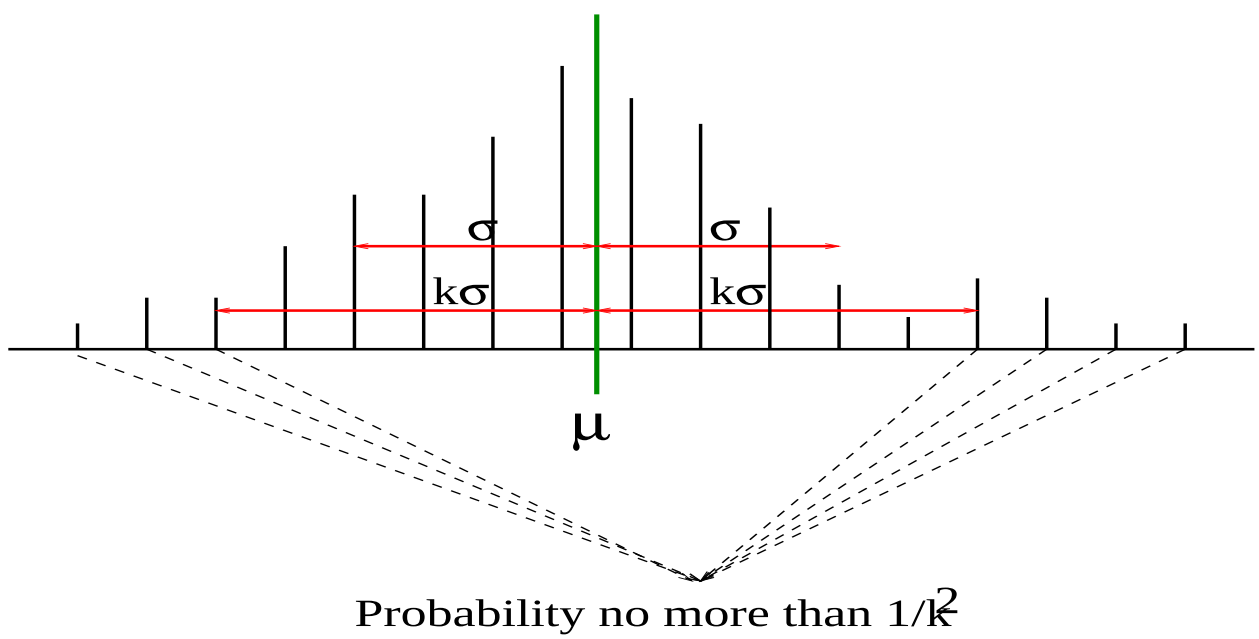
  - Australian males have heights, H with mean $\mu_H = 178cm$ and standard deviation $\sigma_H = 6cm$.
  - Australians have IQs with mean $\mu_{IQ} = 100$ and standard deviation $\sigma_{IQ} = 15$.
  - An Australian male chosen at random is 2m tall, and has an IQ of 83. What is more unusual, his height or his IQ?

$$\begin{aligned} H^* &= (H - \mu_H)/\sigma_H = (200 - 178)/6 = 3.667, \\ IQ^* &= (IQ - \mu_{IQ})/\sigma_{IQ} = (83 - 100)/15 = -1.1333 \end{aligned}$$

Chebychev's Inequality:

$$P\{|X-\mu| \geq k\,\sigma\} \leq \frac{1}{k^2}$$

for <u>any</u> random variable X with mean $E\{X\} = \mu < \infty$ and variance $\sigma^2 < \infty$.



$$\boxed{P\{|X-\mu| \;\geq\; k\,\sigma\} \;\leq\; \tfrac{1}{k^2}}$$

- Recall Markov's inequality: $P\{Y \geq a\} \leq E\{Y\}/a$.
- Set $Y = (X-\mu)^2$, and $a = k^2\sigma^2$.
- Then $\{|X-\mu| \geq k\sigma\} \iff \{Y \geq a\}$.
- Therefore:

$$\begin{aligned} P\{|X-\mu| \;\geq\; k\,\sigma\} &= P\{Y \geq a\} \\ &\leq \frac{E\{Y\}}{a} \\ &= \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}. \end{aligned}$$

$$\boxed{P\{|X^*| \;\geq\; k\} \;\leq\; \tfrac{1}{k^2}, \qquad P\{|X-\mu| \;\geq\; k\} \;\leq\; \tfrac{\sigma^2}{k^2}}$$

Example:

- Recall the less than intelligent Australian, with

$$\begin{aligned} H^* &= (H - \mu_H)/\sigma_H = (200 - 178)/6 = 3.667, \\ IQ^* &= (IQ - \mu_{IQ})/\sigma_{IQ} = (83 - 100)/15 = -1.1333 \end{aligned}$$

- The probability of finding someone that tall, or taller, using $P\{|X^*| \geq k\} \leq \frac{1}{k^2}$, is no more than $1/3.6667^2 = 0.074$.
- The probability of finding someone that unintelligent, or even more lacking in IQ, using the same bound, is no more than $1/(-1.1333)^2 = 0.779$.

- BUT, these are <u>bounds</u> and not approximations. In particular, Chebychev's inequality is not good for k near 1, and useless for $k < 1$.
- BUT, it is the best <u>general</u> inequality that there is!