Mean

$\{x\}$ — dataset

x_i — i'th item

N items

$$\text{mean}(\{x\}) = \frac{1}{N} \sum_{i=1}^{N} x_i$$

- scaling data scales the mean

$$\text{mean}(\{kx\}) = k\,\text{mean}(\{x\})$$

- translating data translates the mean

$$\text{mean}(\{x + c\}) = \text{mean}(\{x\}) + c$$

- $$\sum_{i=1}^{N} (x_i - \text{mean}(\{x\})) = 0$$

Choose μ, such that the sum of squared dist.'s from data points to μ is minimized

$$\mu = \operatorname*{argmin}_a \sum_{i=1}^{N} (x_i - \cancel{\mu}\, a)^2$$

$$\mu = \text{mean}(\{x\}).$$

$$\frac{d}{d\mu} \sum_{i=1}^{N} (x_i - \mu)^2 = 2 \sum_{i=1}^{N} (x_i - \mu) = 0$$

$$\sum_{i=1}^{N} (x_i - \mu) = \sum_{i=1}^{N} x_i - N\mu = 0$$

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i$$

$$\mathrm{std}(\{x\}) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_i - \mathrm{mean}(\{x\})\right)^2}$$

$$= \sqrt{\mathrm{mean}\left(\{(x_i - \mathrm{mean}(\{x\}))^2\}\right)}$$

- translating data does not change std

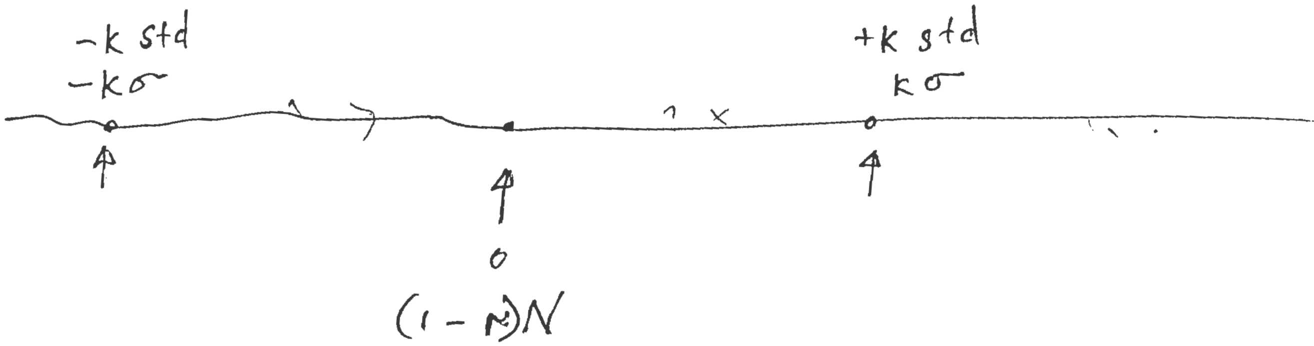
  $\mathrm{std}(\{x+c\}) = \mathrm{std}(\{x\})$

- Scaling data scales std.

  $\mathrm{std}(\{kx\}) = k\,\mathrm{std}(\{x\})$

- for <u>any</u> dataset there are at most $\frac{N}{k^2}$ data items k std away from the mean.

- there is at least 1 data item 1 std away from the mean k std away from mean.



$r\cancel{(1-r)}N$ ← at k std away from

$$\sigma = \sqrt{\frac{(rN)\,k^2\sigma^2}{N}}$$

$$1 = \sqrt{rk^2} \quad \text{so} \quad r = \frac{1}{k^2}$$

≃ 68% within 1 std

≃ 95% within 2 std

≃ 99% within 3std.

$$\mathrm{std}(\{x\})^2 \leqslant \max_i \left(x_i - \mathrm{mean}(\{x\}) \right)^2$$

$$\cancel{N}\,\mathrm{std}(\{x\})^2 = \cancel{\frac{1}{N}}\,\sum_i \left(x_i - \mathrm{mean}(\{x\})\right)^2$$

$$\leqslant \cancel{N} \max_i \left(x_i - \mathrm{mean}(\{x_i\})\right)^2$$

$$\underline{\mathrm{var}(\{x\}) = \mathrm{std}(\{x\})^2}$$

$$\frac{10 \cdot 10^5 + 10^9}{11}$$

$$\text{median}(\{x\}) = \begin{cases} \text{order the data} \\ \text{take the datapoint} \\ \text{in the middle} \\ \text{of ther ordering} \end{cases}$$

Interquartile range:

%

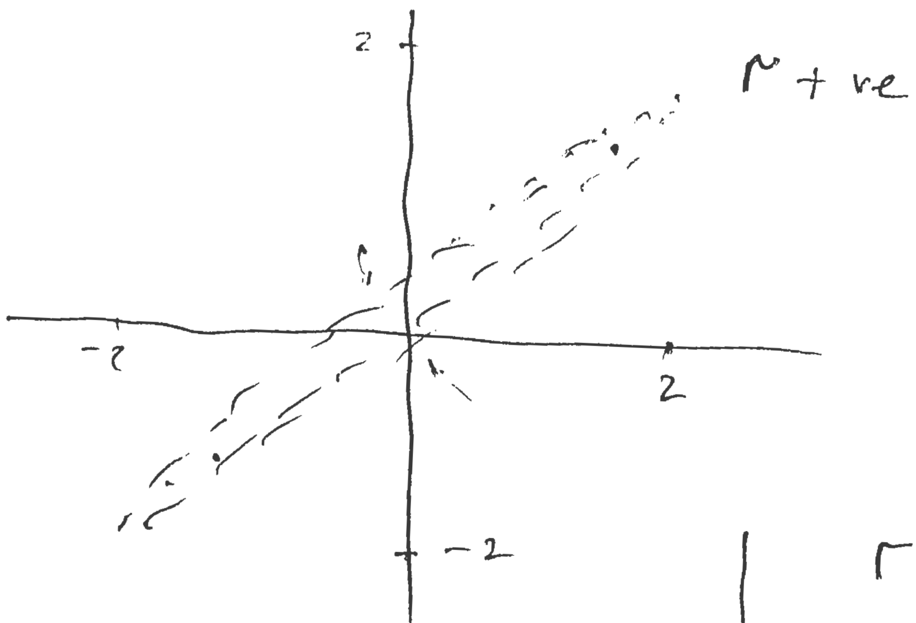
25%ile

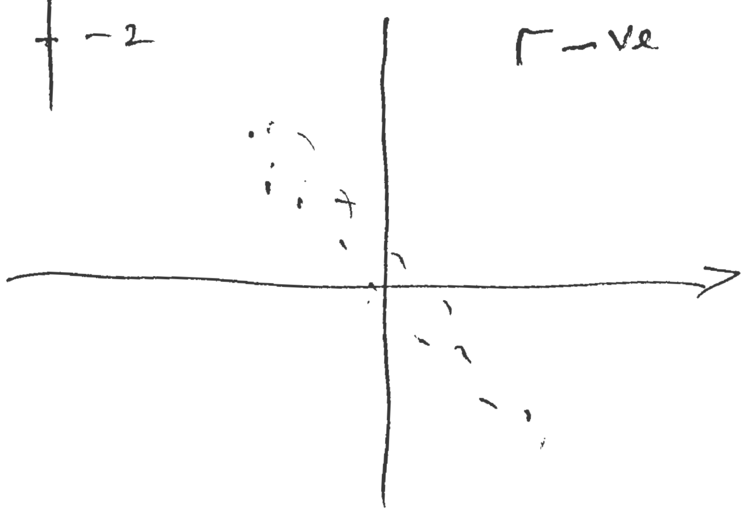
75%ile

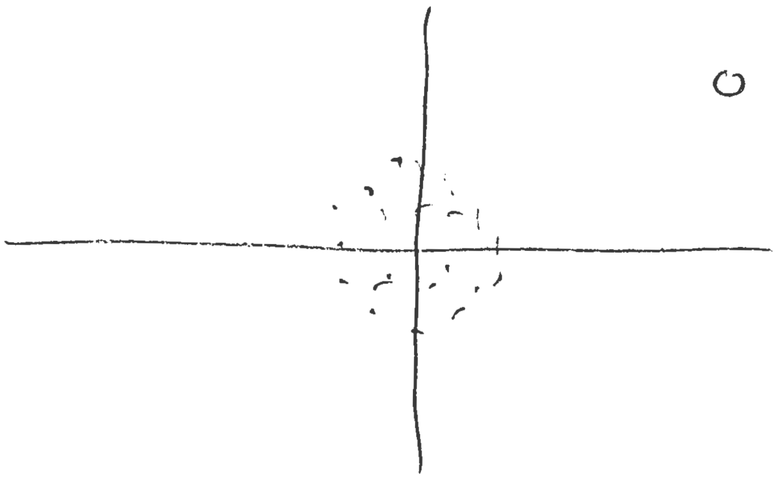
$\text{mean}(\hat{x}) = 0 \qquad \text{Mean}(\hat{y}) = 0$

$\text{mean}(\hat{x}^2) = 1 \qquad \text{mean}(\hat{y}^2) = 1$

$\text{Mean}(\hat{x}\hat{y}) = r =$ correlation coefficient







1) $\mathrm{corr}(x, y) = r = \mathrm{corr}(y, x)$

2) $\mathrm{corr}(\{ax+b, cy+d\})$

$$= \mathrm{sign}(ac)\,\mathrm{corr}(x, y)$$

$-1 \leq \mathrm{corr}(x, y) \leq 1$

$$\mathrm{mean}(\hat{x}\hat{y}) = \frac{1}{N}\sum_{1}^{N} \hat{x}_i \cdot \hat{y}_i = \underline{x} \cdot \underline{y}$$

$$\frac{1}{\sqrt{N}}\begin{pmatrix} \hat{x}_1 \\ \cdot \\ \cdot \\ \cdot \\ \hat{x}_N \end{pmatrix} = \underline{x} \qquad \underline{y} = \frac{1}{\sqrt{N}}\begin{pmatrix} \hat{y}_1 \\ \cdot \\ \cdot \\ \cdot \\ \hat{y}_N \end{pmatrix}$$

$$\underline{x} \cdot \underline{x} = 1 \qquad \underline{y} \cdot \underline{y} = 1$$

$(\hat{x}_i, \hat{y}_i)$, N etc

$(\hat{x}_0, ?)$ $\quad y^p$

$\hat{y}^p = a\hat{x}_0 + b$ $\quad b = 0$

$u_i = \hat{y}_i - \hat{y}_i^p(x_i)$

$\text{mean}(u) = 0$

$$\text{mean}(u) = \text{mean}(\hat{y}_i - a\hat{x}_i - b)$$
$$= \text{mean}(\hat{y}_i) - a\,\text{mean}(\hat{x}_i) - b$$

$\Rightarrow$ $b = 0$

minimize $(\text{var}(u))$

$$\text{var}(u) = \text{var}(\hat{y}_i - a\hat{x}_i)$$
$$= \text{mean}((\hat{y}_i - a\hat{x}_i)^2)$$
$$= \text{mean}(\hat{y}_i^2) - 2a\,\text{mean}(\hat{x}_i\hat{y}_i) + a^2\,\text{mean}(\hat{x}_i^2)$$

r $\quad 1 \quad 1$

$$= 1 - 2ar + a^2$$

$$\frac{d\,\text{var}(u)}{da} = -2r + 2a = 0$$

FACT:

$$\hat{x}_0, ? \rightarrow \hat{x}_0, r x_0$$

$$?, \hat{y}_0 \rightarrow r\hat{y}_0, y_0$$

$?, \hat{y}_0$

$$\hat{x}^p = a y_0 + b$$

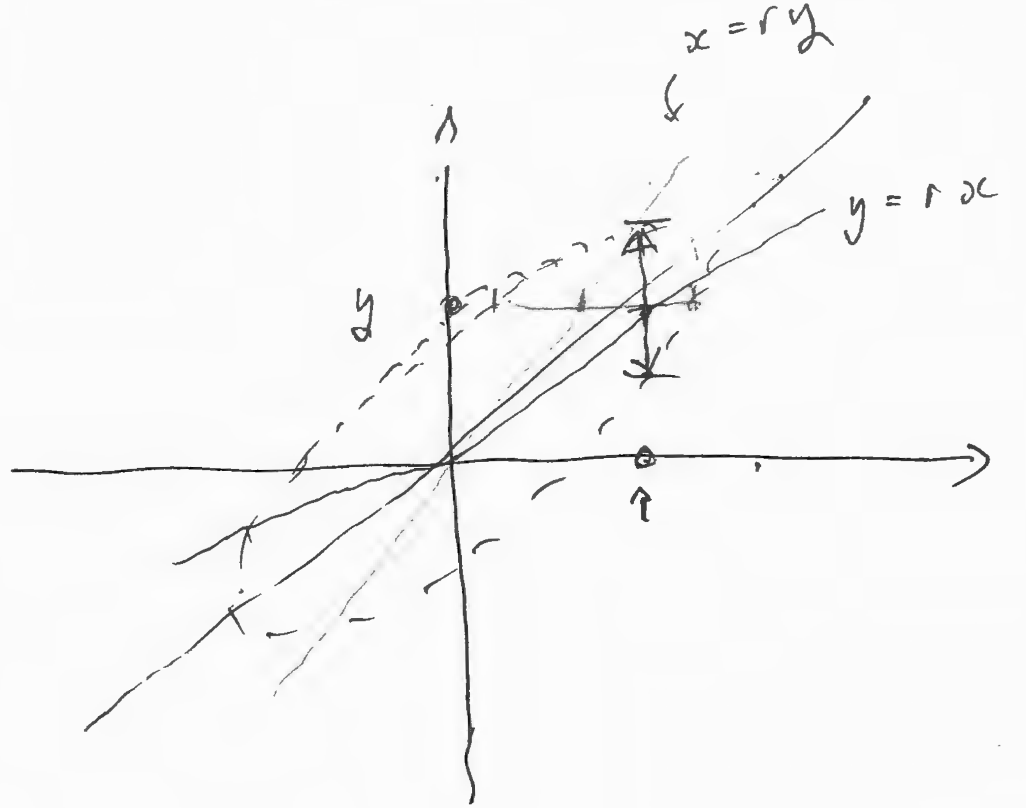
$$u_i = \hat{x}_i - \hat{x}_i^p$$

$$\text{mean}(u) = 0 = \text{mean}(\hat{x}) - a\,\text{mean}(\hat{y}) - b$$

$$\Rightarrow \quad b = 0$$

$$\begin{aligned} \text{var}(u) &= \text{mean}\left((\hat{x}_i - a\hat{y}_i)^2\right) \\ &= \text{mean}(\hat{x}_i^2) - 2a\,\text{mean}(\hat{x}_i \hat{y}_i) \\ &\qquad + a^2\,\text{mean}(\hat{y}_i^2) \\ &= 1 - 2ar + a^2 \end{aligned}$$

$$\rightarrow \quad a = r$$



$u_i = \hat{y}_i - y_i^{\text{AP}}$

$\text{mean}(u_i^2) = \text{var}(u_i) = 1 - 2ar + a^2$

$= 1 - r^2$

$\underset{\sim}{\text{std}(u_i)} = \sqrt{1-r^2}$

1) move to N.C.s

$$\hat{x}_i = \frac{(x_i - mean(x))}{std(x)} \qquad \hat{y}_i = \frac{(y_i - mean(y))}{std(y)}$$

2) get $r = mean(\hat{x}\hat{y})$

3) predict $(\hat{x}_0, \underline{r\hat{x}_0})$

4) goback tonon normalized coords

$$y^p = std(y)\,\hat{y}^p + mean(y)$$

$$= std(y)\, r\, \hat{x}_0 + mean(y)$$

$$= std(y)\, r\, \frac{(x_0 - mean(x))}{std(x)} + mean(y)$$

non

$\{H, T\}$

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

$\{S, \bar{S}\}$

Ω

$\{BG, GB, BBG, BBB,$
$GGB, GGG\}$

$\{GGG, GCG, GGC\}$

$\{CFM, CMF, FCM, \ldots\ldots)$

$\{K, Q, J\}$

$$B^* G^+ B$$

$P(A)$ meaning $\lim_{N \to \infty} \frac{\#(A)}{N}$

$$0 \leq P(A) \leq 1$$

$$\sum_{A_i \in \Omega} P(A_i) = 1$$

$\{1, 2, 3, 4, 5, 6\}$ Ω

$S = \{\text{all even rolls}\}$ $\leftarrow$ Events

$P(S)$

E_1 event

E_2 event

$E_1 \cap E_2$ is event

$E_1 \cup E_2$ is event

E event $\Rightarrow$ $\Omega - E$ event

Σ $\leftarrow$ event space

$\emptyset \in \Sigma$, $\Omega \in \Sigma$

$U \in \Sigma, V \in \Sigma \Rightarrow U \cup V \in \Sigma$

$U \cap V \in \Sigma$

$U \in \Sigma \Rightarrow U^c \in \Sigma$

$$P(X|Y)$$

$$P(Y|x)$$

$$P(y|x) = \frac{P(x|Y)\,P(Y)}{P(X)}$$

$X(\omega)$

$$P(\{\omega : X(\omega) = x\})$$

$$P(x)$$

$$P(\{\omega : X(\omega) \leq x\})$$

$$\{\omega : X(\omega) = x_1\} \cap \{\omega : X(\omega) = x_2\}$$

$$\{\omega : X(\omega) = x_1\} \cap \{\omega : X(\omega) = x_2\}$$

$$\sum_{x \in D} P(\{\omega : X(\omega) = x\}) = 1$$

$$\sum_{x \in D} P(x) = 1$$

1

1	with	p
0		$(1-p)$

$P(H) = p$

$P(0)$	$(1-p)^2$
$P(1)$	$p(1-p)$
$P(2)$	$p(1-p)$
$P(3)$	p^2

q	with	prob p
$-N$	with	$(1-p)$

$X \quad , Y$

$\{\omega : X(\omega) = x\}$ $\qquad\qquad$ $\{\omega : Y(\omega) = y\}$

$P(\{\omega : X(\omega) = x\} \cap \{\omega : Y(\omega) = y\})$

joint probability $\quad P(x, y)$

$$P(\{X = x\} \mid \{Y = y\}) = \frac{P(\{X = x\} \cap \{Y = y\})}{P(\{Y = y\})}$$

$$P(x \mid y) = \frac{P(x, y)}{P(y)}$$

$$P(x \mid y) = \frac{P(y \mid x)\, P(x)}{P(y)}$$

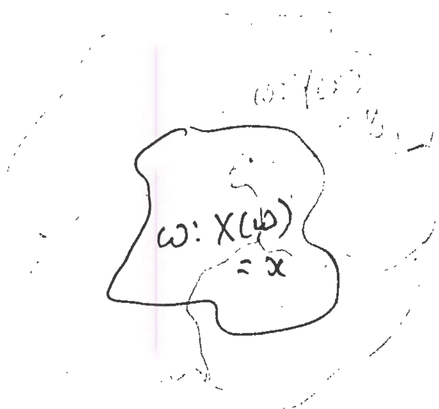
$$\sum_{x \in D} P(x \mid y) = 1$$

A_i st $A_i \cap A_j = \emptyset \quad i \neq j$

<u>and</u> $\bigcup A_i = \Omega$

$$\sum_{y \in D_y} P(x, y) = \sum_{y \in y} P(\{\omega : X(\omega) = x\} \cap \{\omega : Y(\omega) = y\})$$

$= P(x)$



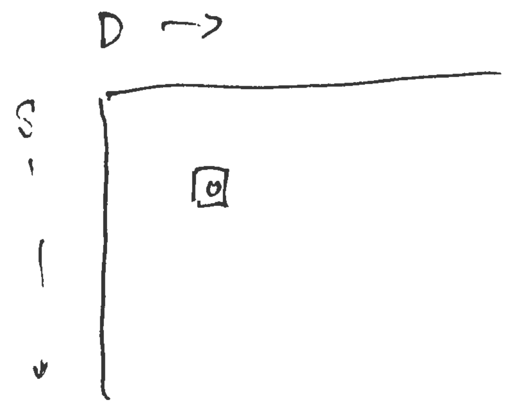
X, Y

$S = X + Y$

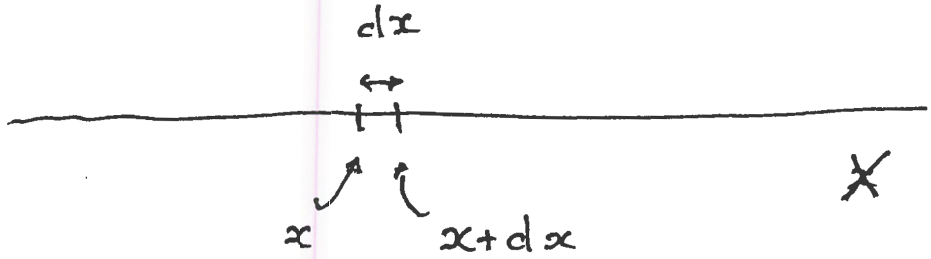
$D = X - Y$

$S = 2$

$S = 3$



$P(x, y) = P(x) P(y)$



$$P(\{\omega : X(\omega) \in [x, x+dx]\}) = p(x)\,dx$$

probability density fn

$$P(\{\omega : X(\omega) \in [a, b]\}) \leftarrow P(\{X = x\}) \leftarrow p(x)$$

$$= \int_a^b p(x)\,dx .$$

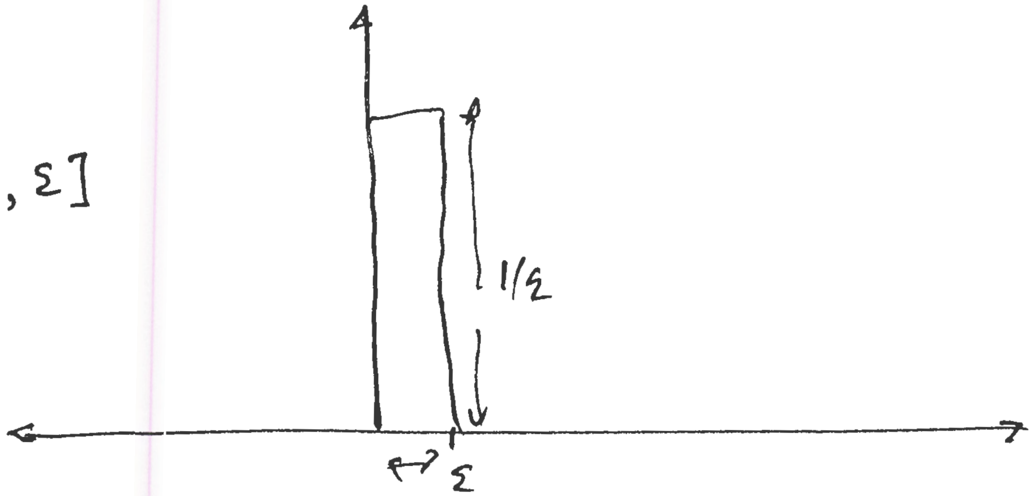
$$\int_{-\infty}^{\infty} p(x)\,dx = 1$$

$$g(x) \geq 0$$

$$\frac{g(x)}{\int_{-\infty}^{\infty} g(x)\,dx} \leftarrow \text{Normalizing const.}$$

$$\frac{\text{the prob that } X \text{ is in } [x, x+dx]}{dx}$$

$[0, \varepsilon]$



$P(H) = p \qquad ; P(T) = 1-p$

Heads, I get q

T, I give r = get $-r$

N times

$$\frac{Npq + (-r)N(1-p)}{N}$$

$$pq + (1-p)(-r)$$

Expectation, Expected value

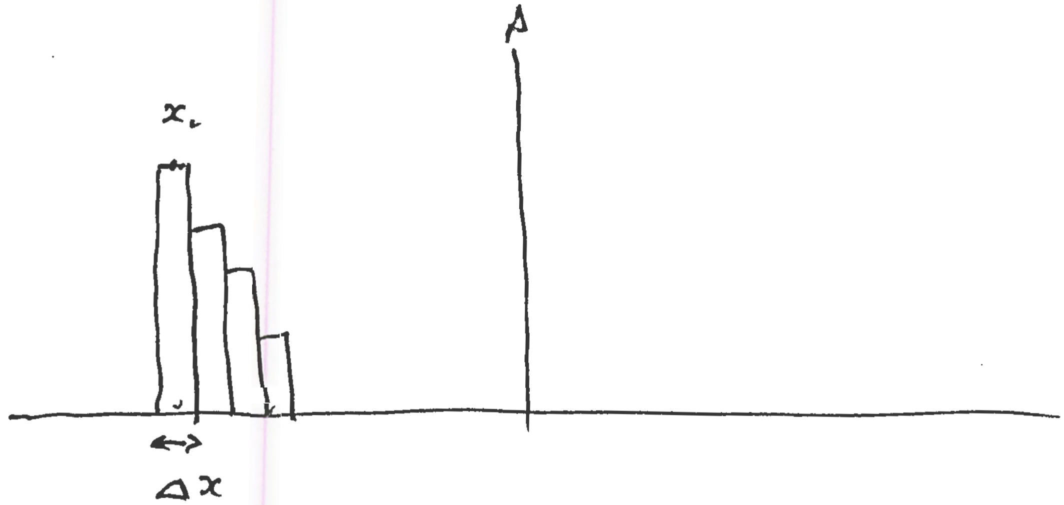
$$\mathbb{E}(X) = \sum_{x \in D} x \, P(\{\omega : X(\omega) = x\})$$

$$f : X \to D_f$$ New RV called F

$$\sum_{u \in D_f} u \, P(\{F = u\})$$

$$\sum_{x \in D} f(x) \, P(\{X = x\}) = \mathbb{E}[f]$$

$p(x)$



$p(x_i)\,\Delta x$

$$\sum_{x_i} x_i \, p(x_i)\,\Delta x$$

$$\int_{-\infty}^{\infty} x\, p(x)\, dx = \mathbb{E}[X]$$

$$\int_{-\infty}^{\infty} f(x)\, p(x)\, dx = \mathbb{E}_p(f)$$

$E[\circ]$

$E[kf] = kE[f]$

$E[f+g] = E[f] + E[g]$

X $\quad E[X] \leftarrow$ mean, expected value

$E[(X - E[X])^2] \leftarrow$ variance

X, Y $\quad E[(X - E[X])(Y - E[Y])] \leftarrow$ covarianc