Sannolikhet och uppräkning.

Sannolikhet representeras med ett värde mellan 0 och 1.
Detta kan även ses som ett procent-tal:
0.63 är alltså 63% sannolikhet.

1 - händer alltid
0 - händer aldrig

0.5 = lika sannolikt som osannolikt

Två sätt att se på sannolikhet: frekventistisk (relativ frekvens) och klassisk sannolikhet.

Relativ frekvens

$$P[A] = \frac{f}{n} = \frac{\text{antalet gånger A inträffade}}{\text{totala antalet observationen}}$$

Exempel:

Vid en kraftanläggning mäts det att under 80 av de närmsta 100 dagarna, slumpmässigt valda från loggarna så nådde kraftförsörjningen ett maximum mellan 18-19. Vad är sannolikheten att max sker mellan 18-19 en annan dag? $\frac{f}{n} = \frac{80}{100} = 0.8$

klassiska formeln:

$$P[A] = \frac{n(A)}{n(S)} = \frac{\text{antalet sätt A kan inträffa}}{\text{antalet utfall experimenten kan ha}}$$

$$S = \{ (\text{brown}, \text{blue}), (\text{blue}, \text{brown}), \underline{(\text{blue}, \text{blue})}, (\text{brown}, \text{brown}) \}$$

Sannolikhet för blåa ögon:

$$P = \frac{1}{4} = 0{,}25$$

En <u>händelse</u> är en delmängd av ett utfallsrum. Alltså
att slumpvariabler antar sina värden.

Tre datorer i månraket:

$$S = \{yyy, yyn, yny, ynn, nyy, nyn, nny, nnn\}$$

En händelse A_1: primära datorsystemet är igång, dvs

$$A_1 = \{yyy, yyn, yny, ynn\}$$

$$A_2 = \{nyy, nyn, nny, nnn\}$$

A_2: primära systemet är nere.

$A_1 = \{yyy, yyn, yny, ynn\}$

$A_2 = \{nyy, nyn, nny, nnn\}$

Händelser är ömsesidigt uteslutande om $A_1 \cap A_2 = \emptyset$

$A_1 \cap A_2 = \emptyset$ är sant för ovanstående!

Faktorial:

$$5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$$

växer enormt fort! Snabbare än expontentialfunktion. Dvs $f(x) = x!$

$f(x) = e^x$

$_nP_r$ - Antalet permutationer av n distinkta objekt r-st åt gången (ordning betydelsefull)

$$_nP_r = \frac{n!}{(n-r)!}$$

$$_9P_4 = \text{"välj 4 av 9"} = \frac{9!}{(9-4)!} = \frac{9 \cdot 8 \cdot 7 \cdot 6 \cdot \cancel{5!}}{\cancel{5!}}$$

$$= 3024$$

Exempel permutation: Utan upprepning!

$_nP_r$ där $n=10$ och $r=1$

upprepning kallas i sannolikhet för "återläggning"

$$\frac{10!}{(10-1!)} = \frac{10!}{9!} = 10$$

=> välj en siffra (av 10 möjliga).

$$_nP_r = {}_{10}P_2 = \frac{10!}{(10-2)!} = \frac{10!}{8!} = \frac{10 \cdot 9}{} = 90$$

dvs för permutation måste objekten vara distinkta

Detta är alltså antalet tal med två siffror som ej är samma!
Vi väljer en siffra på första plats; sedan har vi 9 möjligheter kvar.

Dragning med hänsyn till ordning, utan återläggning = permutation

Dragning med hänsyn till ordning, med återläggning:

Tex antalet tipsrader; 1, x, 2 ; skall gissa 13 matcher

Välj 13 ur 3 möjliga, hänsyn till ordning, med återläggning:

$$n^k = 3^{13} \approx 1,6 \text{ miljoner}$$

Hur många delmängder finns det till en mängd element?

$\Rightarrow 2^n$, Låt varje element "rösta" ja/nej om de skall vara med i delmängden.

Dragning utan hänsyn till ordning och utan återläggning:

$_nC_r$ – "välj r-st av n möjliga"

$\binom{n}{r}$ – binomial koefficient

$$\boxed{(a+b)^n}$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \Rightarrow {}_nP_r = {}_nC_r \cdot r!$$

$$\Rightarrow {}_nC_r = \frac{_nP_r}{r!}.$$

Exempel på kombinationer $\left[{}_nC_r = \binom{n}{r} \right]$:

Hur många ord kan bildas om samtliga bokstäver i ordet "PARALLELL" skall användas?

Vi tänker oss att vi har 9 platser, uppradade från vänster till höger. Välj först fyra platser för "L":en:

$$\binom{9}{4}$$

Välj sedan två 'A':n : $\binom{5}{2}$

Därefter återstående : $3!$

$$\Longrightarrow \binom{9}{4}\binom{5}{2}3! = 7560$$

$$\frac{9!}{4!\,2!}$$

(Dragning med återläggning utan hänsyn till ordning!)

$$\binom{n+r-1}{r}$$

2T6

utfall   individuellt utfall

2 : {1, 1}

3 : {2, 1}, {1, 2}

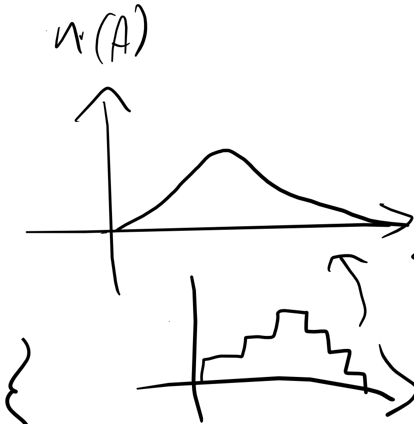
4 : {1, 3}, {2, 2}, {3, 1}

5 : {1, 4}, {2, 3}, {3, 2}, {4, 1}

6 : {1, 5}, {2, 4}, {3, 3}, {4, 2}, {5, 1}

7 : {1, 6}, {2, 5}, {3, 4}, {4, 3}, {5, 2}, {6, 1}

8 : {2, 6}, {3, 5}, {4, 4}, {5, 3}, {2, 6}



T6 → ∞

$n(S) = 36$

$n(A) = 6$

$A \Rightarrow X = 7$

$$\frac{6}{36} = \boxed{\frac{1}{6}}$$

För en tärning:

uniform dist:

$$E[X=x] = \frac{\sum x}{n}$$

$$1+2+3+4+5+6 = 21$$

Expectation (väntevärde)

mean / medel

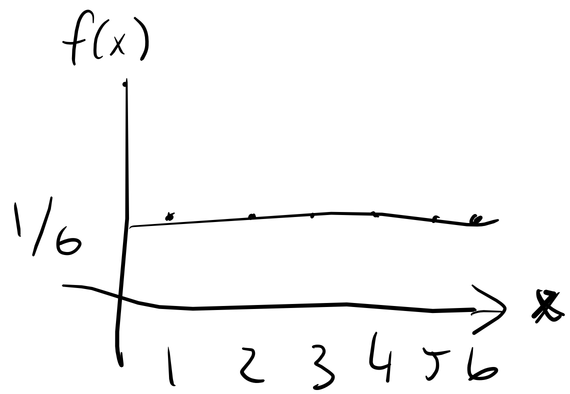
$$\frac{21}{6} = 3.5$$ dvs medlet är 3,5

Detta kallas väntevärde i statistik. $E[X] = \mu$

X är här en slumpvariabel som antar 1–6 med lika (uniform) sannolikhet.

$P[X \leq 6]$ = hela distributionen

$E[X] = 3.5$



$$E[X] = \sum_x H(x) f(x)$$

uniform dist. $H(x) = x$

$f(x)$ = täthetsfunktion (sannolikhet för varje värde).

$S(X) = \{1, 2, 3, 4, 5, 6\}$ dvs $x \in X$ tar värden 1–6.

$$\sum_x x f(x) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + \dots + 6 \cdot \frac{1}{6}$$

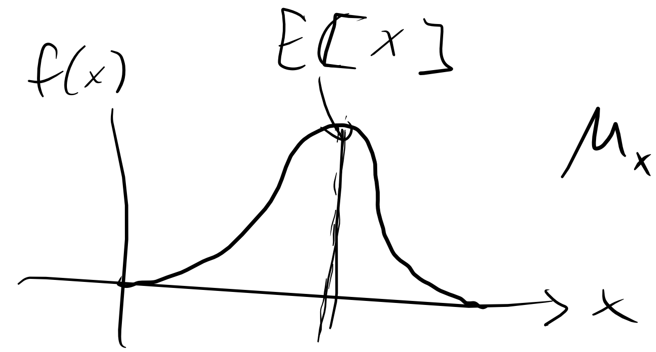
$$= \frac{1+2+3+4+5+6}{6} = 3,5$$

Istället T8 (8-sidig tärning): $\sum_{i=1}^{8} \frac{1}{8} \cdot i = \frac{1+2+3+4+5+6+7+8}{8}$

$$= \frac{36}{8} = 4,5$$

Väntevärde för X (med utfall x):

$$E[X] = \sum_x x f(x)$$



$$\mu_x = 70$$

Ett till exempel:

En hjärtmedecin för efterbehandling vid hjärtinfarkt undersöks. Låt X vara antalet hjärtslag per minut och patient.

x	40	60	68	70	72	80	100
$f(x)$	.01	.04	.05	.80	.05	.04	.01

$$\sum_x f(x) = 1$$ stämmer!

$$E[X] = 40(.01) + 60(.04) + 68(.05) + 70(.8) + 72(.05) + 80(.04) + 100(.01)$$
$$= 70$$

Väntevärdet är ett av flera viktiga <u>kvantitavia</u> mått.
De kallas moment (till en distribution).

Ett annat är <u>varians</u>.

Låt X vara en s.v. med medel μ. Variansen av X $\operatorname{Var} X$, eller σ^2, ges av

$$\operatorname{Var} X = \sigma^2 = E[(X-\mu)^2]$$
$$= E[X^2] - (E[X])^2$$

(formel för beräkning)

Exempel (medicin exemplet)

x	40	60	68	70	72	80	100
$f(x)$	.01	.04	.05	.80	.05	.04	.01

$$E[x^2] = \sum_x x^2 f(x) = 40^2(.01) + 60^2(.04) + \dots + (100^2)(.01)$$
$$= 4926.4$$

$$\text{Var } X: \quad E[x^2] - (E[x])^2 = 4926.4 - 70^2$$
$$= 26.4$$

Svårt att tolka! Inga enheter och svårt att tyda.

Standardavvikelsen (Standard Deviation):

$$\sigma = \sqrt{\mathrm{Var}\,X} = \sqrt{\sigma^2}$$

$\sqrt{26.4} = \underline{5.14}$ i hjärtslag!

Stickprovs avvikelsen (standard error) i allmänna fall;

går mot 0 när antalet datapunkter (utfall) går mot ∞.