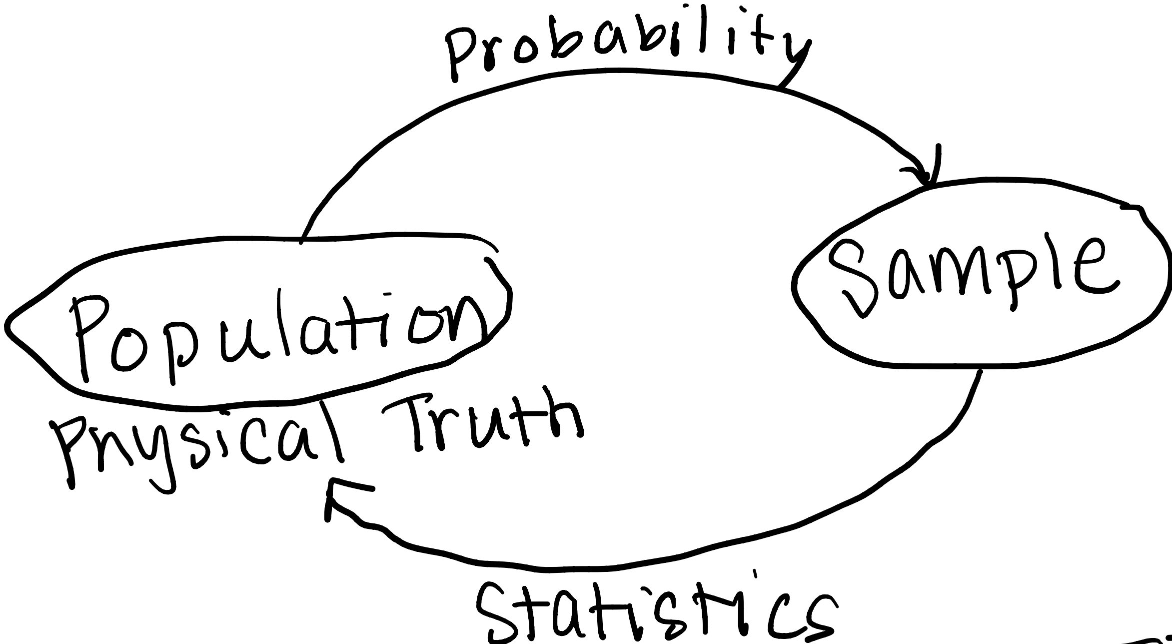
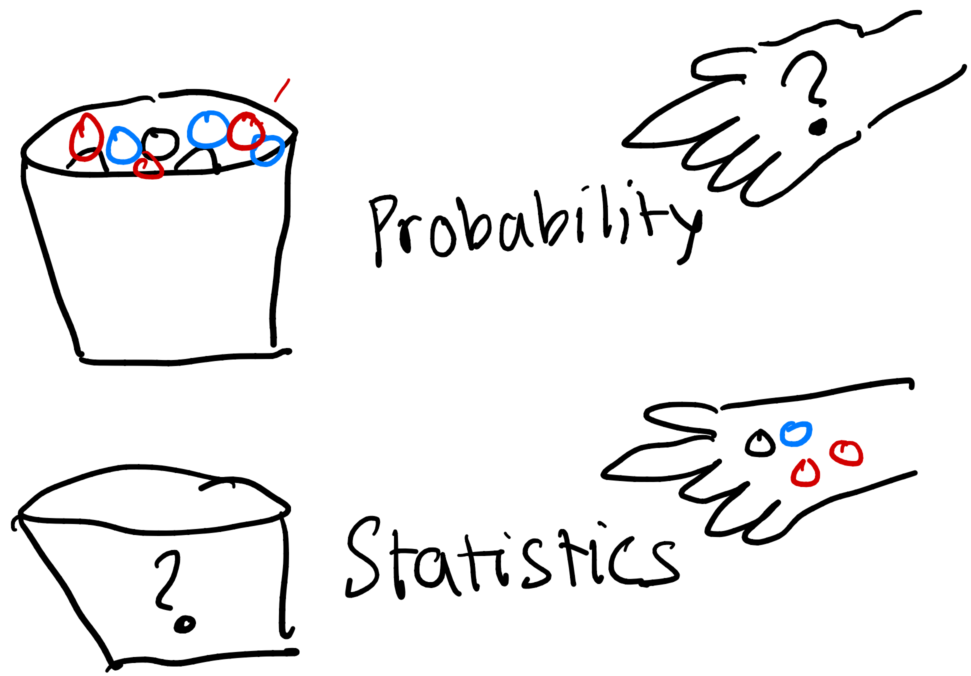
Introduction to Statistics: Probability

Sarah Mancina

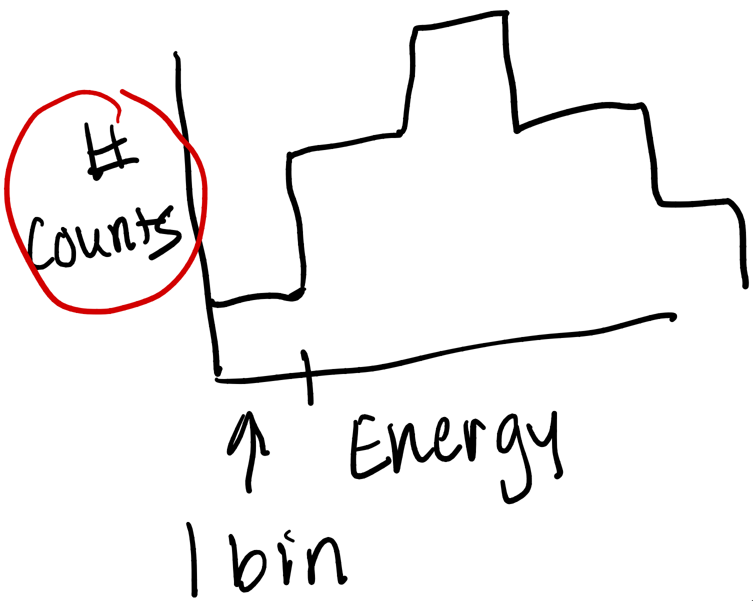
6/15/2020





Quatatative

- Discrete
- Continuous

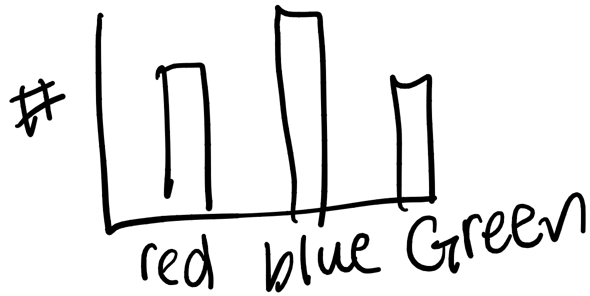


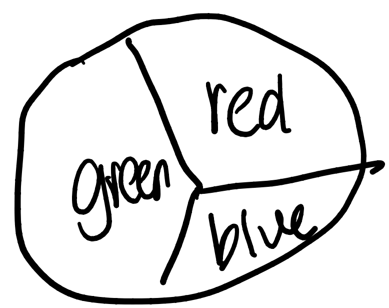
frequency: $f_i = \dfrac{n_i}{\sum_j^{bins} n_j}$

density: $d_i = \dfrac{n_i}{w_i \sum_j^{bin} n_j}$

Qualitative

- Nominal
- Binary
- Ordinal

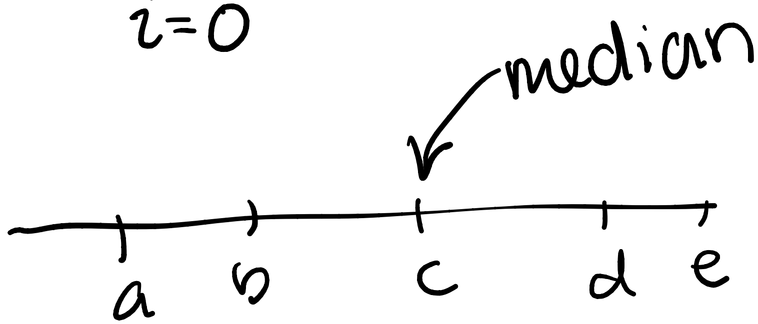




Descriptive Statistics

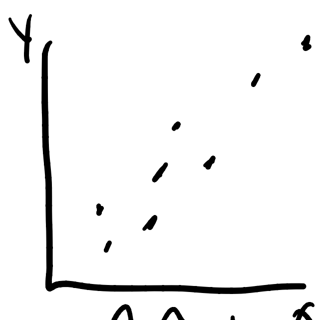
mean: $$\bar{x} = \sum_{i=0}^{N} \frac{1}{n} x_i$$

median:



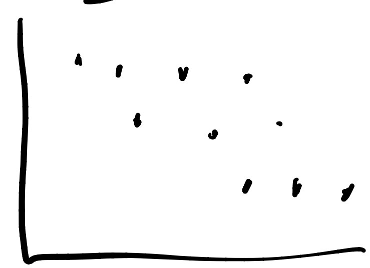
std. deviation: $$\sigma = \sqrt{\sum_{i=0}^{N} \frac{1}{n}(x_i - \bar{x})^2}$$

correlation: $$\rho = \frac{\sum_{i=0}^{N} \left[(x_i - \bar{x})(y_i - \bar{y})\right]}{\sigma_x \sigma_y}$$



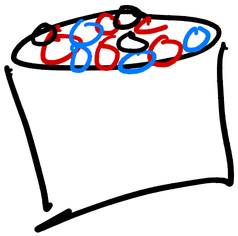
$\rho \simeq 1$

$\rho = 0$



$\rho \simeq -1$

$\rho \simeq 0$



30r
60b
20y

Given Set A, the complent, $\bar{A}$, set $\forall a \in S$, $a \notin A$

Given Set A and B, $A \cup B$, set $\forall a \in S$, $a \in A$ <u>or</u> $a \in B$

Given Set A and B, $A \cap B$, set $\forall a \in S$, $a \in A$ <u>and</u> $a \in B$

$P(E \in [10, 20)\ \text{TeV})$

Kolmogorov Axioms

1. $\forall A$ in S, $P(A) \geq 0$
2. $P(A) + P(\bar{A}) = 1$
3. $A_1, A_2, A_3 \ldots$ are mutually exclusive, $P(A_1 \cup A_2 \ldots) = \sum_{i}^{N} P(A_i)$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

$$P(A \cap B) = P(B|A)P(A) = P(A|B)P(B)$$

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Bayes' Theorem

powerlaw: $\phi(E) \propto \left(\frac{E}{100\,\mathrm{TeV}}\right)^{-\gamma}$? $P(\gamma|E_i)$? ?

known!

$\Rightarrow P(E_i|\gamma) = L(\gamma|E_i)$?

Statistics

Frequintist Probability

$$P(x) = \lim_{N \to \infty} \frac{n}{N}$$

# of times x occured

Trials

Discrete

Probability Mass Function

$$\sum_{i=0}^{max} PMF(x_i) = 1$$

$$CMF(r) = \sum_{i=0}^{r} PMF(x_i)$$

$$\overline{g(x)} = \sum_{i=0}^{max} g(x_i) PMF(x_i)$$

$g(x) = x$

$$\overline{x} = \sum_{i=0}^{max} x_i PMF(x_i)$$

Continuous

Probability Density Function

$$\int_{min}^{max} PDF(x)dx = 1$$

$$CDF(r) = \int_{min}^{r} PDF(x)dx$$

$$\overline{g(x)} = \int_{min}^{max} g(x) PDF(x)dx$$

$$\overline{x} = \int_{min}^{max} x PDF(x)dx$$

Binomial Distribution

$P(1) = 1/6 \quad P(\text{not } 1) = 5/6$

$P(1 \,\&\, 1 \,\&\, \text{not} 1 \,\&\, \text{not} 1 \,\&\, \text{not} 1) = 1/6 \times 1/6 \times 5/6 \times 5/6 \times 5/6$

$P(\text{Roll 1 twice}) = \binom{5}{2}(1/6)^2(5/6)^3$

$\binom{5}{2} = \frac{5!}{2!3!}$

$P(K) = \binom{n}{K} p^K (1-p)^{(n-K)}$

n = total # of trials
K = # of sucesses
p = probability of K for 1 trial

$\bar{K} = n \cdot p \qquad \sigma = \sqrt{np(1-P)}$

Poisson Distribution

$$P(K) = \frac{e^{-\lambda} \lambda^K}{K!} \qquad \bar{K} = \lambda$$

$$\sigma = \sqrt{\lambda}$$

On average, 10 HE ν in IC per year

- Whats the probability, 15 ν in a given year?

$\lambda = 10$

$$P(15) = \frac{e^{-10} 10^{15}}{15!} = 0.03472$$

- Whats the prob, 1 HE ν in a month?

$\lambda = \frac{10}{12} = 0.8\overline{333}$

$$P(1) = \frac{e^{-0.833}(0.833)^1}{1!} = 0.3622$$

Normal Distribution

Central Limit Theorem

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

$\bar{x} = \mu$

$\sigma = \sigma$