

LMS Statistics and Hypothesis Testing

November 2023

LMS statistics and hypothesis testing

Chapter 1. Random events and probability theory

- i) What is probability
- ii) Discrete distributions
- iii) Continuous distributions

Chapter 2. Confidence intervals and central limit theorem

- i) Confidence intervals
- ii) Z and t distribution
- iii) The Central limit theorem

Chapter 3. Hypothesis testing

- i) Formulate hypothesis H_0, H_1
- ii) Quantify significance. p-value
- iii) Comparing means, t-test.

1.1 Random events and probability

* Random events ("stochastic")

Something whose output we don't know

* Probability: number $\in [0, 1]$ quantifying certainty / "surprise".

Example: tossing coins (H, T)

$P(H) = 0 \rightarrow$ certain I will never get H

$P(H) = 1 \rightarrow$ always get H

$0 < P(H) < 1 \rightarrow$ level of uncertainty / "surprise"

* Unitarity: The sum of probabilities for all possible outcomes x_i must add up to 1.

$$\sum_{\forall x_i} P(x_i) = 1$$

Example: tossing coins

$$P(H) + P(T) = \frac{1}{2} + \frac{1}{2} = 1 \quad \checkmark$$

Example: rolling dice

$$P(1) + P(2) + P(3) + P(4) + P(5) + P(6) = \frac{1}{6} + \frac{1}{6} + \dots + \frac{1}{6} = 1 \quad \checkmark$$

1.2 Discrete probability distributions

* Discrete: number of possible outcomes is a finite number. } dice

i) Binomial distribution

"How many times X I get a specific result in n trials,

if the probability of each success is $p"$

$$\left\{ B(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x} \right. \quad \left. \begin{array}{l} x: \text{number of successes} \\ n: \text{number of trials} \\ p: \text{probability each success} \end{array} \right.$$

Example: Probability of 5 times H tossing 10 times a coin

$$B(5; 10, \frac{1}{2}) = \binom{10}{5} \left(\frac{1}{2}\right)^5 \left(1-\frac{1}{2}\right)^{(10-5)} = 0.246$$

✖
Exercise

Example: Probability of 3 times a 6 in 10 dice

$$B(3; 10, \frac{1}{6}) = \binom{10}{3} \left(\frac{1}{6}\right)^3 \left(1-\frac{1}{6}\right)^{(10-3)} = 0.155$$

✖
Exercise

Example: Probability of passing exam (A, B, C) of 10 quest.

$$B(5; 10, \frac{1}{3}) = \binom{10}{5} \left(\frac{1}{3}\right)^5 \left(1-\frac{1}{3}\right)^{(10-5)} = 0.137$$

✖
Exercise

iii) Poisson distribution

"How many times I observe an event in an interval, know λ "

$$\left\{ \begin{array}{l} P(x; \lambda) = \frac{e^{-\lambda} \lambda^x}{x!} \\ \quad \quad \quad \left\{ \begin{array}{l} x: \text{number of times I observe} \\ \lambda: \text{observed average.} \end{array} \right. \end{array} \right.$$

Example: Probability of observing 3 new patients, with $\lambda = 5$.

$$P(3; 5) = \frac{e^{-5} 5^3}{3!} = 0.140$$

~~Excel~~

Example: Probability of observing 5 or less patients in same λ .

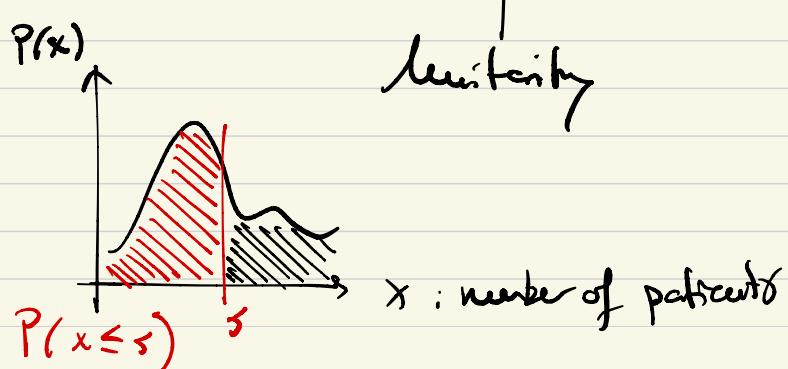
$$\begin{aligned} P(x \leq 5; 5) &= P(1) + P(2) + \dots + P(5) \\ &= \sum_{x=1}^5 P(x; \lambda=5) = 0.616 \end{aligned}$$

~~Excel~~

Accumulative distribution function cdf

Example: Probability of more than 5

$$P(x > 5, \lambda=5) = 1 - cdf(5) = 0.392$$



1.3 Continuous distributions

* Continuous: amount of possible outcomes is infinite / uncountable

Case 5: 2 outcomes (H, T) $\rightarrow P = \frac{1}{2}$ Discrete cases "Mass dist."

Dice: 6 outcomes ($1, 2, 3, 4, 5, 6$) $\rightarrow P = \frac{1}{6}$ Frequentist definition /

Continuous variable ($T, h, \text{conc.}$): ∞ outcomes $\rightarrow P = \frac{1}{\infty} = 0$ WTF

↳ Frequentist approach does not work

Need a new mathematical object "Density distribution"

i) Discrete case

Probability $P \in [0, 1]$

Unitarity $\sum_{x_i} P(x_i) = 1$

ii) Continuous case

Density $f(x)$

Unitarity $\int_{-\infty}^{\infty} dx f(x) = 1$

* Gaussian

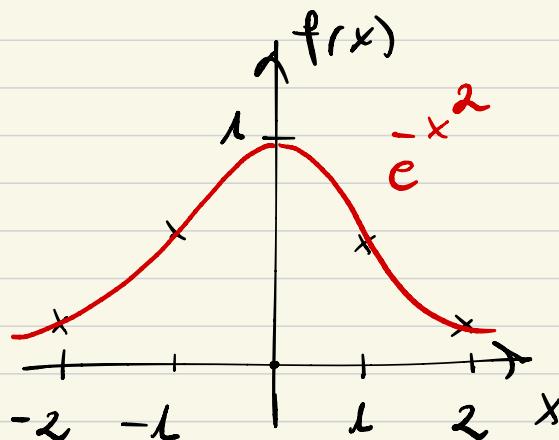
i) Gaussian distribution

"Random variable x centered at some μ and spreaded $\sigma"$

$$\left\{ f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \right. \left. \begin{array}{l} \mu: \text{mean value} \\ \sigma: \text{standard deviation} \end{array} \right\}$$

* Consider simplest case $\mu=0$; $\sigma=L$; ignore normalization factor

$$f(x) = e^{-x^2} ; \text{ plot same values}$$



$$x = 0; f(0) = e^{-0} = 1$$

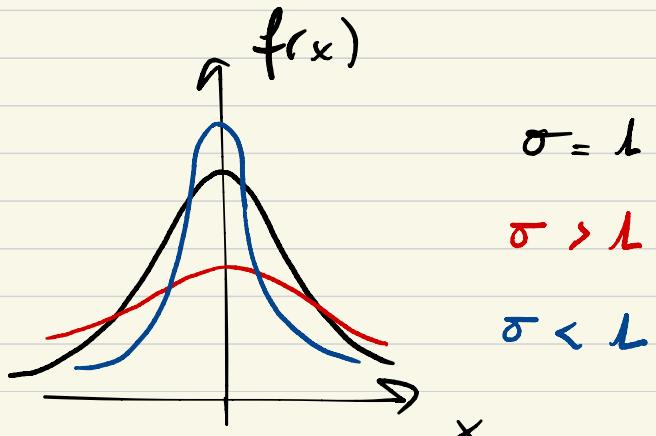
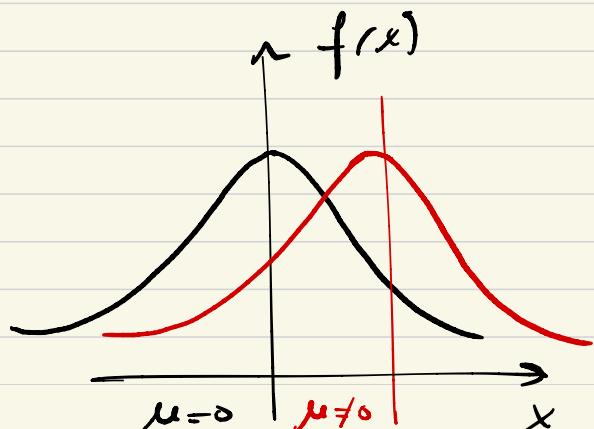
$$x = L; f(L) = e^{-L^2} = \frac{1}{e} = 0.37$$

$$x = -L; f(-L) = e^{-L^2} = \frac{1}{e} = 0.37$$

$$x = \pm 2; f(\pm 2) = e^{-4} = \frac{1}{e^4} = 0.018$$

* Switch back on μ and σ $\left. \begin{array}{l} \mu \text{ will displace the central value} \\ \sigma \text{ will drive sharper/wider shape} \end{array} \right\}$

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



* Add normalization factor

i) Remember discrete probability: Normality $\sum_{x_i} P(x_i) = 1$

ii) Continuous case. Density distributions

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

*
Exercise

$$\int_{-\infty}^{+\infty} dx f(x) = 1$$

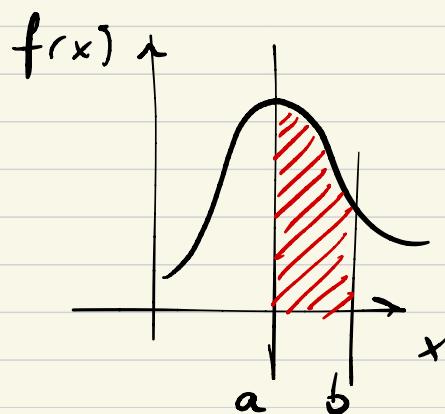
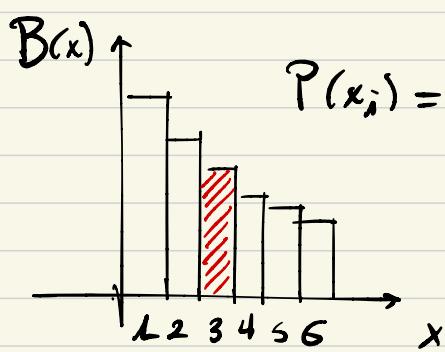
* for $f(x)$ to behave as a probability, we add a normalization factor

$$\left\{ f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \right\} \text{ Gaussian distribution; } \int_{-\infty}^{\infty} dx f(x) = 1$$

* Particular case $\mu=0; \sigma=1$

$$\left\{ f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \right\} \text{ Normal distribution } N(0,1)$$

* In continuous distributions, we can only compute probability in a given range. Never a single value.

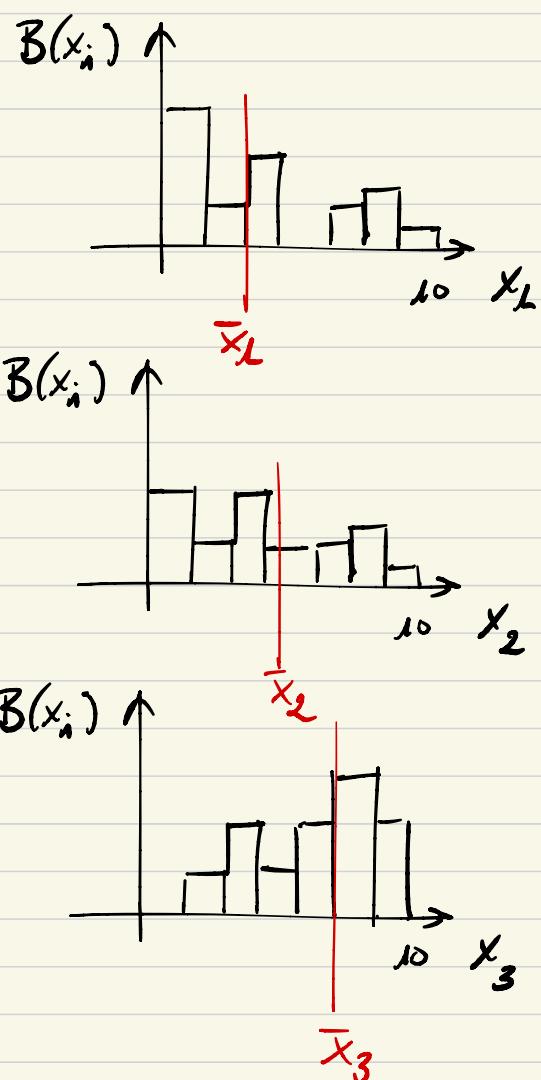


$$P(a < x < b) = \text{cdf}(b) - \text{cdf}(a)$$

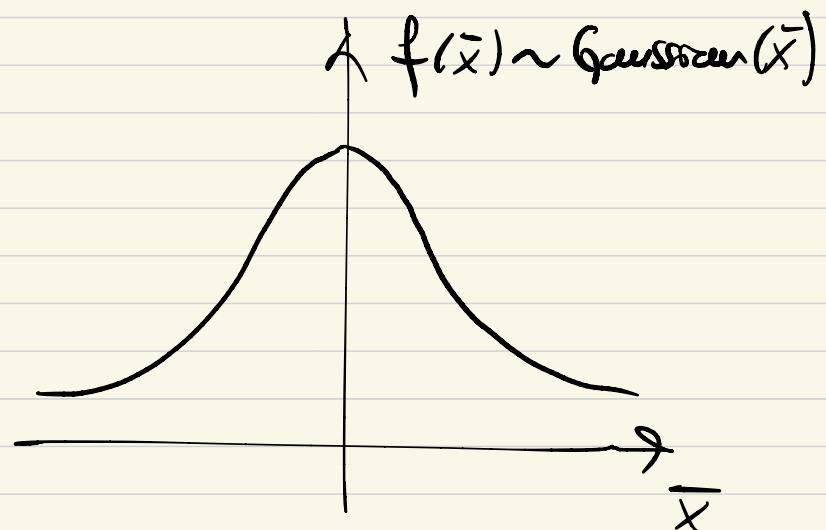
2.1 The Central Limit Theorem

"For a set of n independent and identically distributed (iid) random variables x_i , the sample means tends to a gaussian distribution for large values of n , regardless of how x_i are themselves distributed"

- Consider 10 rolls of a dice. How many times I observe a 6. Binomial distributed $B(x; 10, 1/6)$. Repeat N times.



* As the sample size n increases, the distribution of means \bar{x}_i tends to a gaussian distribution



* Mean value μ of the \bar{x}_i variable
* Standard error (SE)

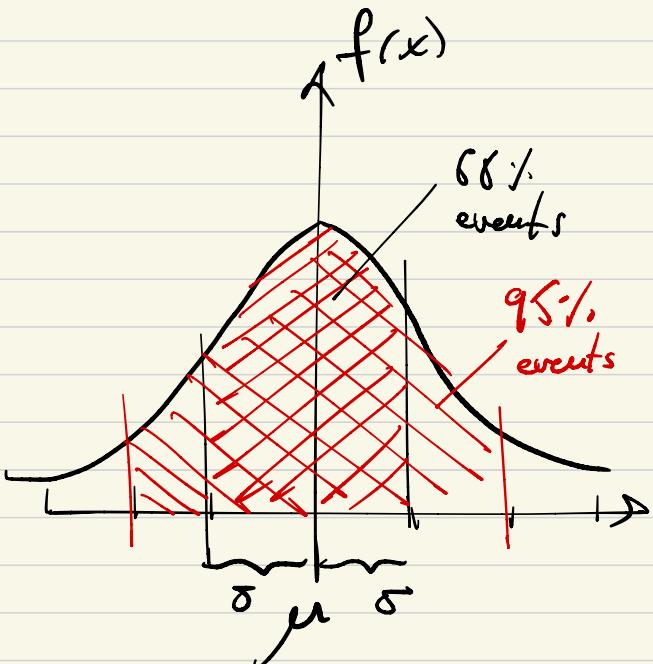
- Build new random variable

$$\bar{x} = \{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_N\}$$

$$SE = \frac{\sigma_i}{\sqrt{n}} ; \text{ as } n \text{ increase, } \sqrt{n} \downarrow SE$$

2.2 Confidence Intervals.

i) Consider a Gaussian distribution $f(x; \mu, \sigma)$



- * 68% of events are contained in the $(\mu - \sigma, \mu + \sigma)$ interval. 1σ conf. int.
- * 95% of events are contained in the $(\mu - 2\sigma, \mu + 2\sigma)$ interval. 2σ conf. int.
- * 99% of events are contained in the $(\mu - 3\sigma, \mu + 3\sigma)$ interval. 3σ conf. int.

ii) Any random variable can be normalized / "standardized" just by subtracting the mean and dividing by the standard deviation.

x : random distributed; mean μ, σ .

standardized x ; } $z = \frac{x-\mu}{\sigma}$ { follows a $N(0,1)$ distribution,
also referred to as z score or z distribution.

iii) We want to estimate the mean μ
The best estimator is the sample mean \bar{x} , or "average"
Normally, we want to build

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} ; \text{ such that } P(\bar{x} - z < \mu < \bar{x} + z) = P_{\text{critical value}}$$

③ Hypothesis testing

i) Check if an observation is compatible with a given hypothesis H_0

* statistic test : a function / number computed out of our data
(t-test, χ^2 test, Fisher test, Wald test, ...)

* p-value : probability of obtaining a result at least as extreme as the one actually observed, under the assumption that our null hypothesis H_0 was correct.

(how likely / unlikely was to observe this result)

} i) \uparrow p-value : it was likely to obtain this result. **Accept H_0 .**
ii) \downarrow p-value : // unlikely // . **Reject H_0 .**

ii) General approach

1. Formulate null hypothesis H_0 and significance level α .
2. Make measurement / observations.
3. Compute statistic test
4. Compute p-value
5. If $p\text{-value} < \alpha$, reject H_0 . Otherwise, accept H_0 .

iii) A bit of history

Karl Pearson (1857-1936). Pearson's "p" values for χ^2 test.

Mathematical statistics.

Ronald Fisher (1890-1962) "Statistical methods for research" (1925)

Example: Check if a die is biased.

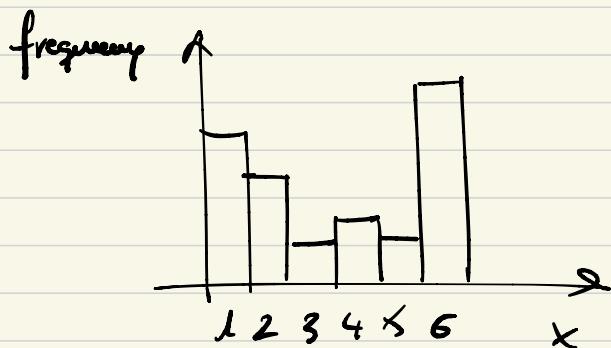
* You suspect your opponent is using a loaded die

i) Formulate null hypothesis $H_0: p = \frac{1}{6}$

ii) alternative ii $H_A: p > \frac{1}{6}$

Choose significance level $\alpha = 0.01$

iii) Collect data. Roll die 100 times



43 times out of 100
we get a 6.

iv) Compute probability of 43 or less 6's, given H_0 .

Given by Binomial distribution with $p = 1/6$

$$B(43; 100, \frac{1}{6}) \sim L_{\text{calf}}$$

Exercice

v) Compute p-value. Probability of obtaining at least a value as extreme as 43.

$$\text{p-value} = 1 - B_{\text{calf}}(43; 100, \frac{1}{6}) = 1^{-10}$$

↪ p-value $< \alpha$. → evidence to reject H_0 .

Example: check if coin is fair

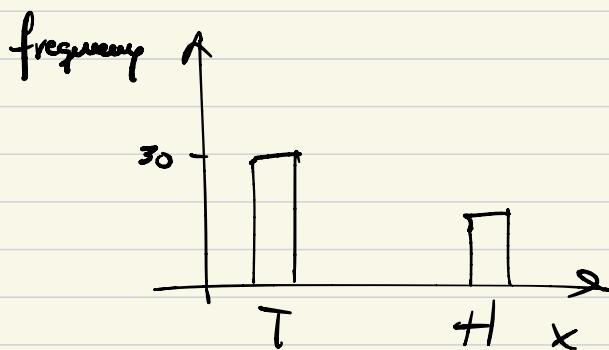
* You suspect your opponent is using a biased coin

i) Formulate null hypothesis $H_0: p = \frac{1}{2}$

ii) alternative " $H_1: p \neq \frac{1}{2}$

Choose significance level $\alpha = 0.05$

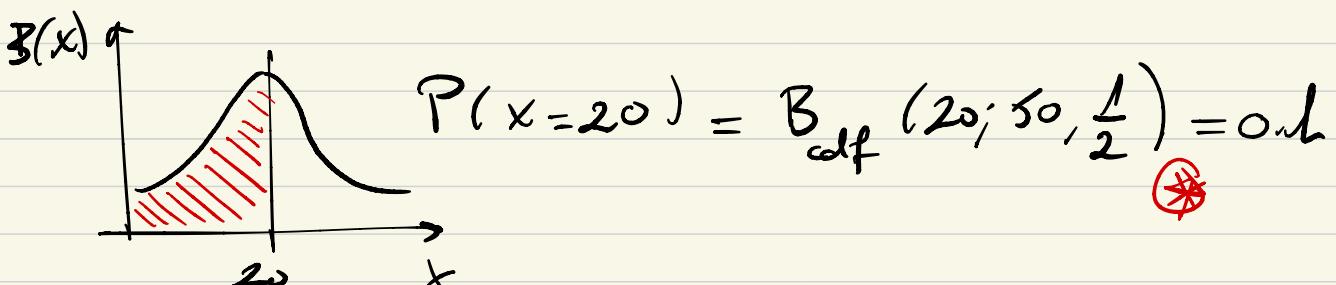
iii) Collect data. Roll die 100 times



20 heads in 50
flips of a coin

iv) Compute probability of 43 or less 6's, given H_0 .

Given by Binomial distribution with $p = 1/6$



*

v) Compute p-value.

Now our H_1 is just $p \neq \frac{1}{2}$, either larger or smaller.

$p\text{-value} = 2 \cdot \text{cdf}(20) = 0.2 < \alpha$ \nexists evidence
to reject H_0 .

