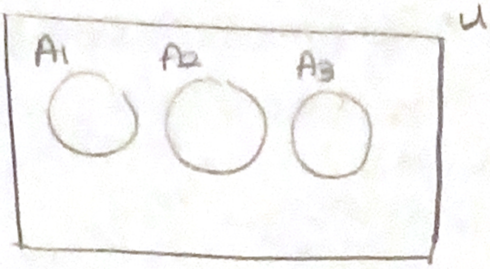
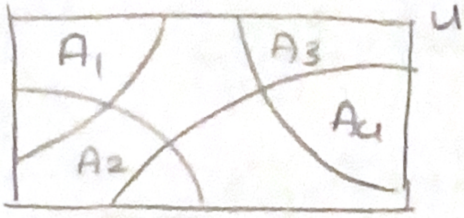
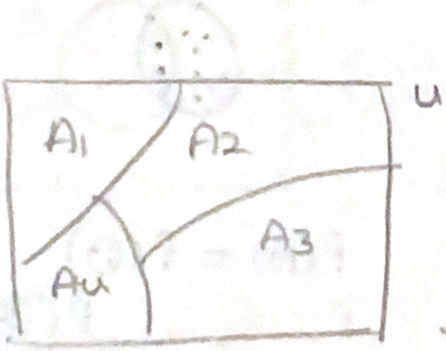
Reintroducing Probability



- Mutually exclusive ✓
- Collectively exhaustive ✗



- Mutually exclusive ✗
- Collectively exhaustive ✗

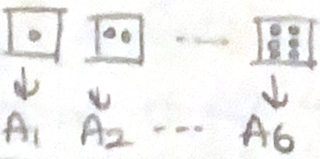


- mutually exclusive ✓
- Collectively exhaustive ✓

→ Partition

→ Biri olurken diğeri olamaz
Sadece bunlar olabilir

Fine-grained partitions → Sample Space



→ ME ✓
CE ✓ → Fine grained ✓ (1, 2, 3, 4, 5, 6)

$$\underbrace{A_1 \quad A_2}_{B_1} \quad \underbrace{\cdots \quad A_6}_{B_2} \rightarrow \text{Fine grained} \times$$

Given $\{A_1, A_2 \cdots A_3\}$ is a sample space

$A_i \rightarrow P(A_i) \rightarrow$ Event'in gerçekleşme olasılığı

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

$\sum P(A_i) = 1$

$P(U) = 1$

$P(\emptyset) = 0$

$X = A_i \cup A_j \cup \cdots A_4$

↳ Any outcome should be combination of events

Ex

outcome is a prime number

Conditional Probability

- $P(A|B)$ → A given B
  - $P(A|B)$ ↓ event I'm looking for
  - B ↳ event that happened

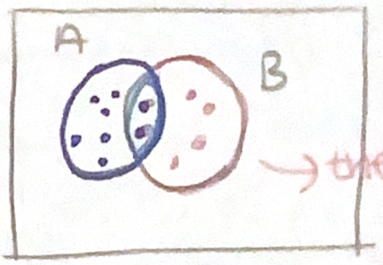
- $P(AB) = P(B) \cdot P(A|B)$
  - $P(AB)$ ↓ how can A & B happen together?
  - $P(B)$ ↳ let B happen
  - $P(A|B)$ → given B has happened let A happen

- $P(A|B) = \dfrac{P(AB)}{P(B)}$
- $P(B|A) = \dfrac{P(AB)}{P(A)}$

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

Bayes' Rule

→ ☆

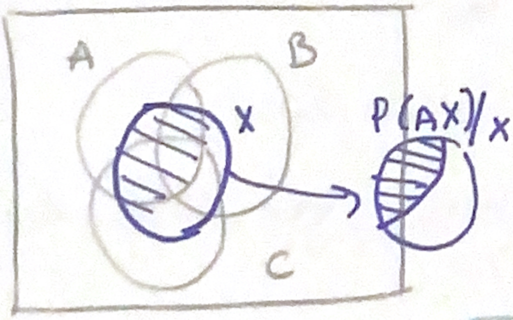


→ the universal set is this now

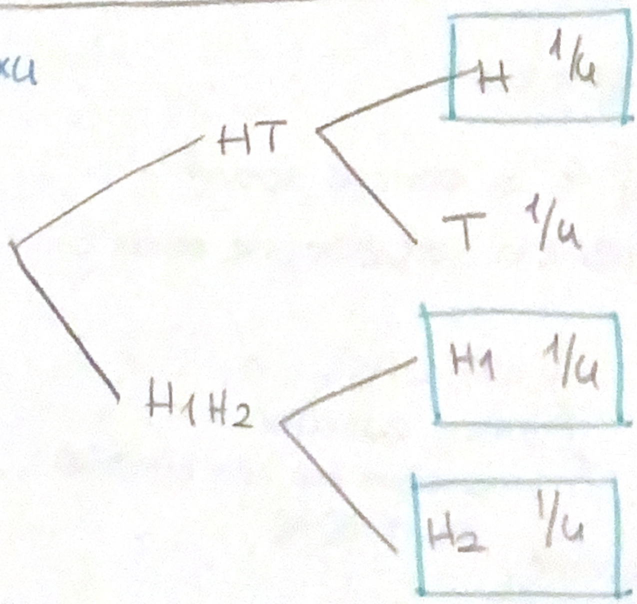
$P(A) = P(☺)$

$P(A|B) = \dfrac{P(☺)}{P(☺)}$ → intersection $P(AB)$

what is prob of A & B happening together when sample space is B happening?



exu



Given: Heads is observed

$P(\text{Heads observed}) = P(H_1) + P(H_2) + P(H)$

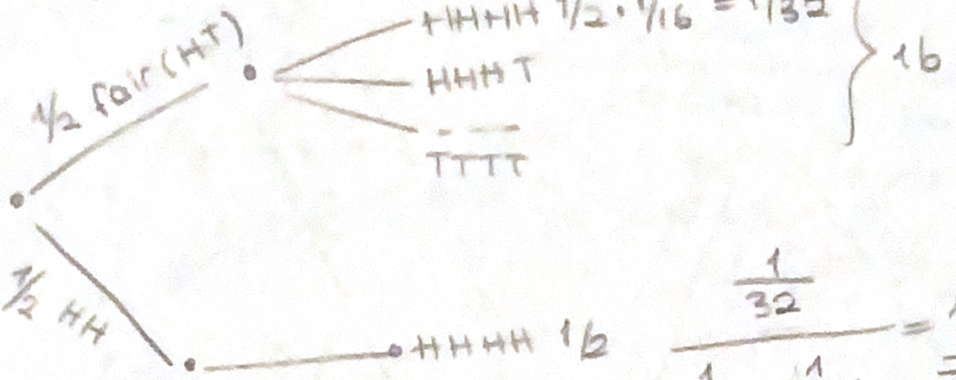
$P(H_1) = \dfrac{P(H_1 H_o)}{P(H_o)}$

New sample space

HW1

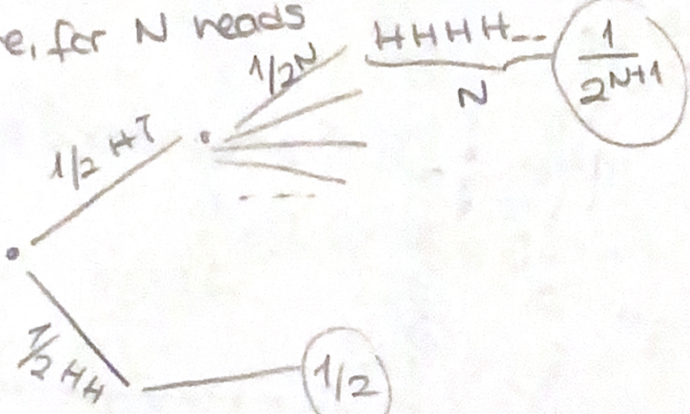
1) Two coins in pocket. One is fair, other heads on both sides. Pick one, toss randomly, get HHHH.

i) Probability that the coin was fair? P(fair)?



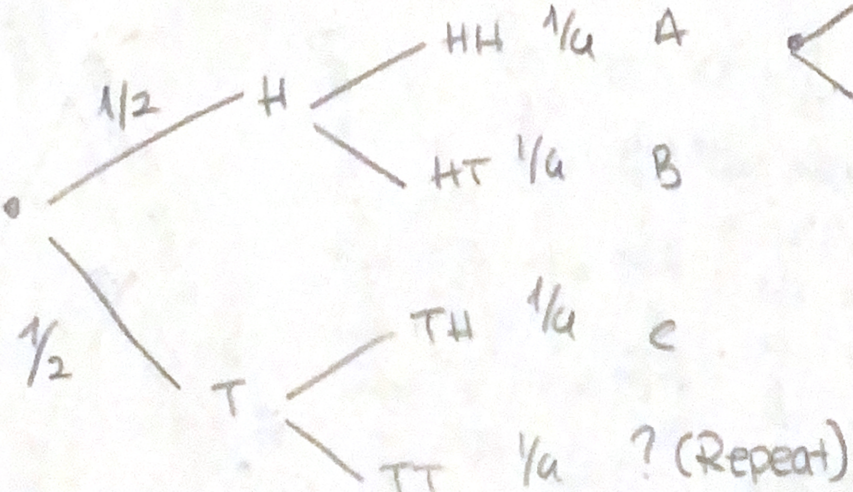
$$\frac{\frac{1}{32}}{\frac{1}{32}+\frac{1}{2}} \;(16) = \underline{\underline{1/17}}$$

ii) Same, for N heads

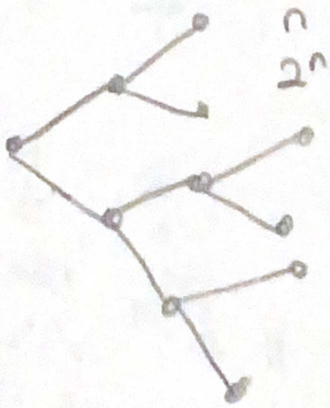


$$\frac{\frac{1}{2^{N+1}}}{\frac{1}{2^{N+1}}+\frac{1}{2}} \;(2^{N+1}) = \underline{\underline{\frac{1}{2^{N+1}}}}$$

2) Three Choices with a coin I have a fair coin. Make a procedure to pick one of three choices (A, B, C) so each choice is equally likely.



↓
not



n flips
2^n events in SS

2^n equally likely outcomes

$\frac{2^n}{3}$ is not possible

• Approximate to 1/3

32-bit number (32 flips)

2^{32} ← Sample space size

Statistical Inference*

part I

*making deductions on a value in a population out of a sample

$\theta \rightarrow$ Parameter of interest
Can be mean, variance or a future value if you're predicting.
ex: average weight of all newborn babies 👶

point estimation

Try to predict θ (a single value) from a random sample.

- $\hat{\theta} = \theta + \text{error}$
  - $\hat{\theta}$: our guess
  - θ: real value

Unbiased estimators $\rightarrow E(\hat{\theta}) = \theta + 0$ • $\text{Bias} = E(\hat{\theta}) - \theta$

Unbiased Estimators

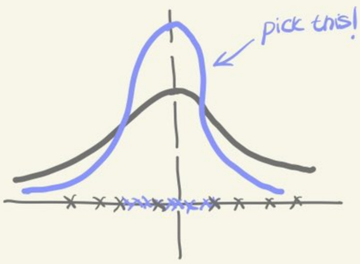
θ (real)	$\hat{\theta}$ (our guess)
μ (population mean)	$\bar{X}$ (sample mean)
σ^2 (population variance)	s^2 (sample variance)

$s^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}$ — random samples; n: sample size; why -1? if we use n it becomes biased. it underestimates σ^2

- Standard error of $\hat{\theta}$ $= \sigma_{\hat{\theta}} \rightarrow$ deviation between θ & $\hat{\theta}$

→ If you have two point estimators, pick the unbiased one.

→ If both are unbiased, pick the one with smallest variance.

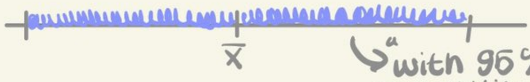


More detail:
- Maximum Likelihood Estimation
- Method of Moments

confidence intervals

Try to predict an interval which contains θ with a confidence level.

Assumptions: Population is normal. Data is random.



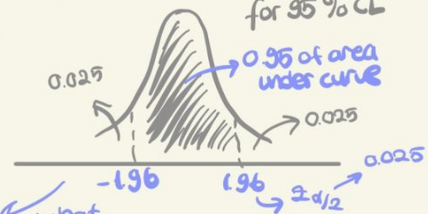
"with 95% probability the population mean (μ) is in this interval"

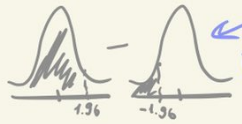
- α = significance level
- $(1-\alpha)$ = confidence level

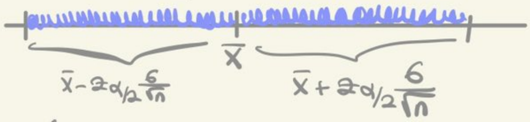
(these are determined by you!)

- z-score = "how many standard deviation is a point away from mean?"

$z_{\alpha/2} \rightarrow$ for 95% CL





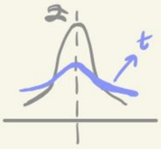


- $P\Big(\underbrace{\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}}_{\text{interval}}\Big) = 1 - \alpha$

→ Wider interval → less predictive power

→ For smaller samples → use t value. → fatter twin of z

$t_{\alpha/2, n-1}$ — this is also called "v"

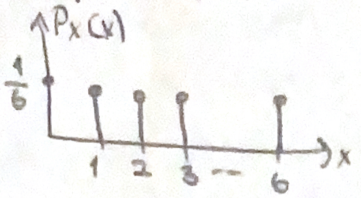


@mervenoyann

Discrete RVs

$P(x)=P(X=x)$ (pmf)

x	0	1	2 ...
$P(x)$	0.2	0.3	0.2



Cumulative Distribution Function

$F(x)=P(X\le x)=\sum_y p(y)$ (at most x)

$F(x\le 2)=0.7$ ← at most 2

$P(a\le X\le b)=F(b)-F(a-)$ ← a-1

ex: $P(2\le x\le b)=F(5)-F(1)$

Expected Value

$E[X]=\sum x\,P_X(x)=\mu_X$

$E[h(x)]=\sum h(x)p(x)$

Variance

$V(X)=E[X^2]-E[X]^2$

Joint Distribution

$P(x,y)=P(X=x\ \&\ Y=y)$

$\sum_x \sum_y P(x,y)=1$

$P(x,y)$		0 · 1 ...
x	5	
	10	→ $P(10,1)$ → sum all values on row to get $X=x$

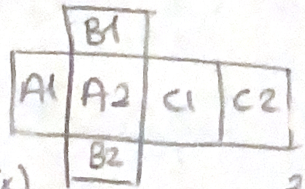
$P_X(x)=\sum_y P(x,y)$

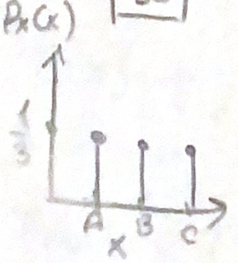
condition for independence ↓

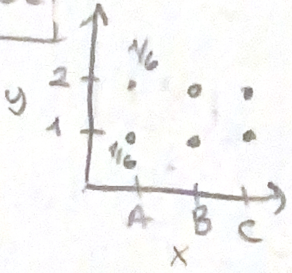
$P_{X,Y}(x,y)=P_X(x)\cdot P_Y(y)$

$X=A\ B\ C$

$Y=1,2$





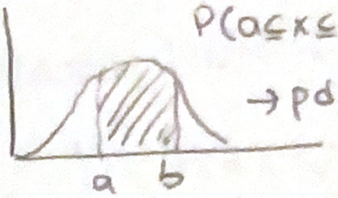


$E(XY)=\sum_x\sum_y xy\,P(X=x,Y=y)$

Continuous RVs

$P(a\le x\le b)=\int_a^b f(x)\,dx$

→ pdf



$P(x=0)\rightarrow P(a\le x\le b)=P(a<x<b)$
$=P(a\le x<b)=P(a<x\le b)$

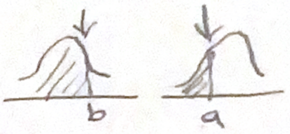
Cumulative Distribution Function

$F(x)=P(X\le x)=\int_{-\infty}^{x} f(y)\,dy$

$P(X>a)=1-F(a)$

$P(a\le x\le b)=F(b)-F(a)$

$\text{PDF} \underset{\text{integrate}}{\overset{\text{derive}}{\leftrightarrows}} \text{CDF}$



Expected Value

$\mu_X=E(X)=\int_{-\infty}^{\infty} x\,f(x)\,dx$

$E[h(x)]=\mu_{h(x)}=\int_{-\infty}^{\infty} h(x)f(x)\,dx$

Variance

$V(X)=E[X^2]-E[X]^2$

Joint Distribution

$P(X,Y)=\iint f(x,y)\,dx\,dy$

$P(a\le X\le b,\ c\le Y\le d)=\int_a^b\int_c^d f(x,y)\,dy\,dx$

$f_X(x)=\int_{-\infty}^{\infty} f(x,y)\,dy$ ↳ integrate over y

↓ marginal of x

$f_Y(y)=\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$

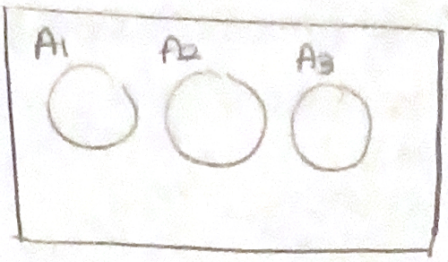
Condition for Independence

$f_{X,Y}(x,y)=f_X(x)\,f_Y(y)$

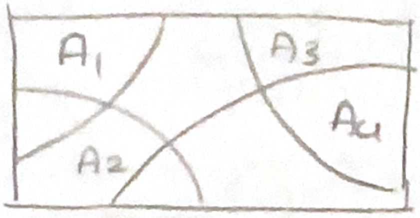
$E[g(X,Y)]=\iint \underbrace{g(x,y)}_{x\cdot y} f(x,y)\,dx\,dy$

$E[XY] \longrightarrow$

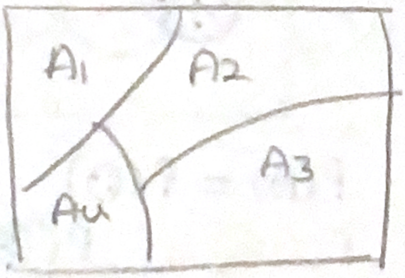
Reintroducing Probability



- Mutually exclusive ✓
- Collectively exhaustive ×



- Mutually exclusive ×
- Collectively exhaustive ×

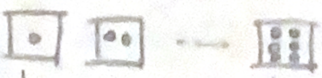


- mutually exclusive ✓
- Collectively exhaustive ✓

→ Partition

→ Biri olurken diğeri olamaz
Sadece bunlar olabilir

Fine-grained partitions → Sample Space



→ ME ✓ CE ✓ Fine grained ✓ (1, 2, 3, 4, 5, 6)

$A_1 \quad A_2 \; \cdots \; A_6$

$\underbrace{A_1 \quad A_2}_{B_1} \; \underbrace{\cdots \; A_6}_{B_2}$ → Fine grained ×

Given $\{A_1, A_2 \cdots A_3\}$ is a sample space

$A_i \rightarrow P(A_i) \rightarrow$ Event'in gerçekleşme olasılığı

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

$\sum P(A_i) = 1$

$P(u) = 1$

$P(\emptyset) = 0$

$X = A_i \cup A_j \cup \cdots A_4$

↳ Any outcome should be combination of events

__Ex__

outcome is a prime number

Conditional Probability

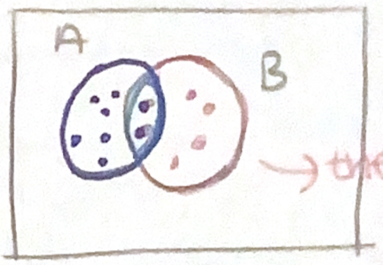
- $P(A|B) \rightarrow$ A given B
  - A ↓ event I'm looking for
  - B ↳ event that happened

- $P(AB) = P(B) \cdot P(A|B)$
  - $P(AB)$ ↓ how can A & B happen together?
  - $P(B)$ ↳ let B happen
  - $P(A|B)$ ↗ given B has happened let A happen

- $P(A|B) = \dfrac{P(AB)}{P(B)}$
- $P(B|A) = \dfrac{P(AB)}{P(A)}$

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

Bayes' Rule

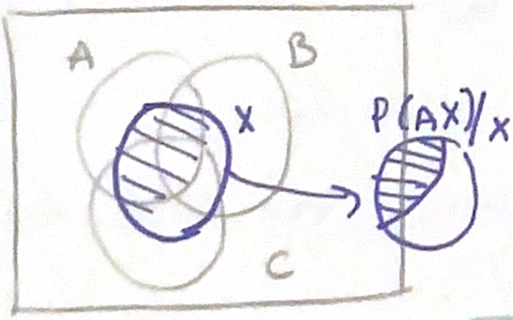
→ ☆ 

→ the universal set is this now

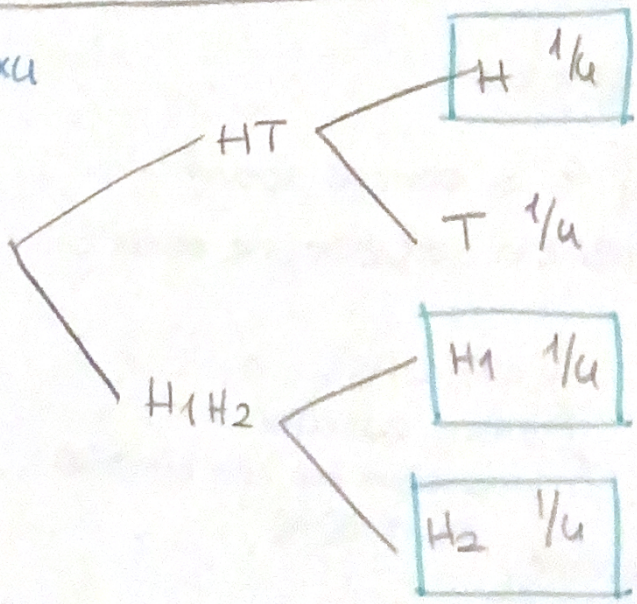
$P(A) = P(\odot)$

$P(A|B) = \dfrac{P(\odot)}{P(\odot)}$ ↗ intersection $P(AB)$

What is prob of A & B happening together when sample space is B happening?



exu



Given: Heads is observed

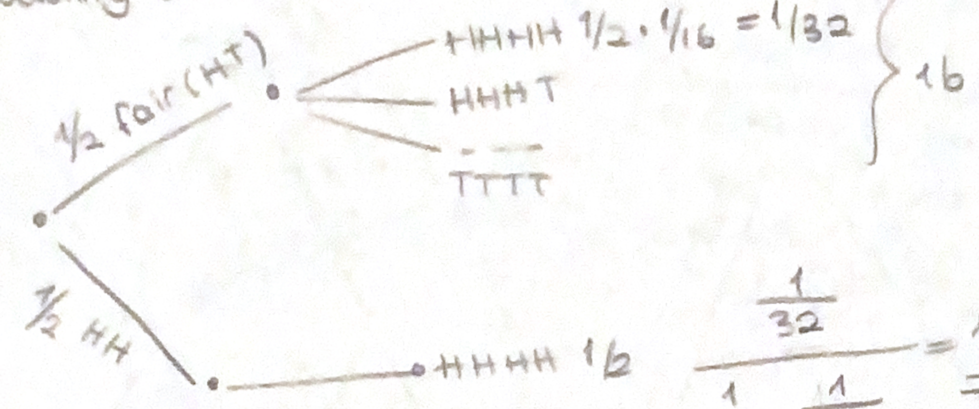
$P(\text{Heads observed}) = P(H_1) + P(H_2) + P(H)$

$P(H_1) = \dfrac{P(H_1 H_o)}{P(H_o)}$ ← New sample space

HW1

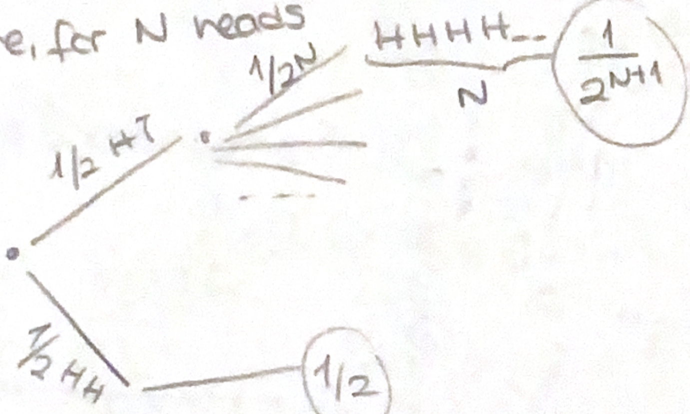
1) Two coins in pocket. One is fair, other heads on both sides. Pick one, toss randomly, get HHHH.

i) Probability that the coin was fair? P(fair)?



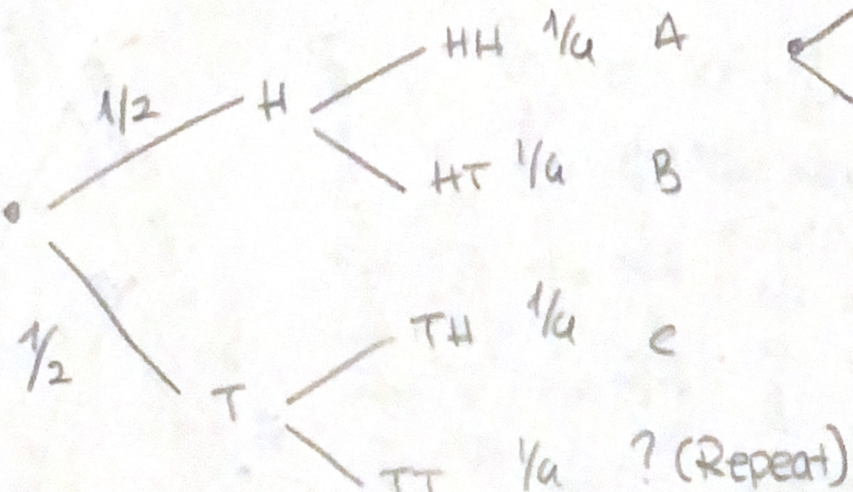
$$\frac{\frac{1}{32}}{\frac{1}{32}+\frac{1}{2}}_{(16)} = \underline{\underline{1/17}}$$

ii) Same, for N heads

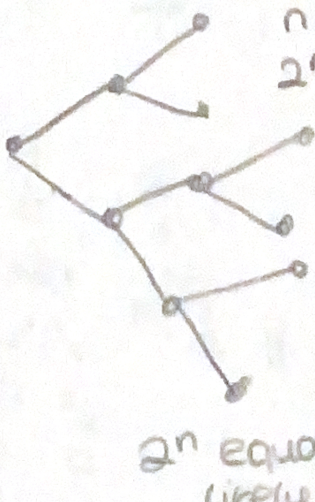


$$\frac{\frac{1}{2^{N+1}}}{\frac{1}{2^{N+1}}+\frac{1}{2}}_{(2^{N+1})} = \underline{\underline{\frac{1}{2^{N+1}}}}$$

2) Three Choices with a coin I have a fair coin. Make a procedure to pick one of three choices (A, B, C) so each choice is equally likely.



↓ not



n flips
2^n events in SS

2^n equally likely outcomes

$\frac{2^n}{3}$ is not possible

- Approximate to 1/3

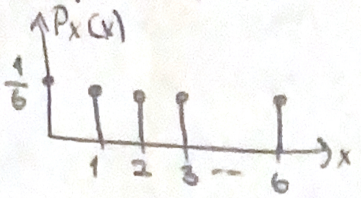
32-bit number (32 flips)

2^{32} ← Sample space size

Discrete RVs

$P(x) = P(X = x)$ (pmf)

x	0	1	2 ...
$P(x)$	0.2	0.3	0.2



Cumulative Distribution Function

$F(x) = P(X \le x) = \sum_y p(y)$ (at most x)

$F(x \le 2) = 0.7$ ← at most 2

$P(a \le X \le b) = F(b) - F(a-)$

exa $P(2 \le x \le 5) = F(5) - F(1)$ ← a−1

Expected Value

$E[X] = \sum x P_X(x) = \mu_X$

$E[h(x)] = \sum h(x) p(x)$

Variance

$V(X) = E[X^2] - E[X]^2$

Joint Distribution

$P(x, y) = P(X = x \ \& \ Y = y)$

$\sum_x \sum_y P(x, y) = 1$

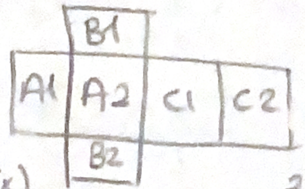
$P(x,y)$		y: 0 · 1 ...
x	5	↓
	10	→ $P(10,1)$ →

sum all values on row to get $X = x$

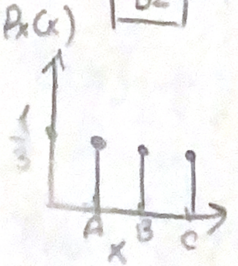
$P_X(x) = \sum_y P(x, y)$

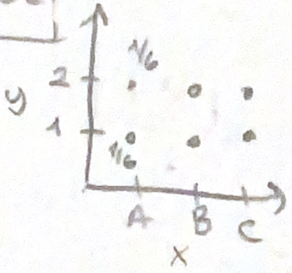
condition for independence

$P_{X,Y}(x,y) = P_X(x) \cdot P_Y(y)$



$X = A\,B\,C$
$Y = 1, 2$

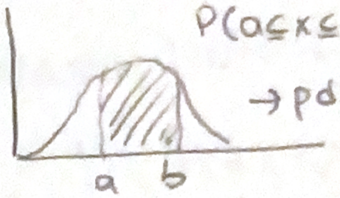




$E(XY) = \sum_x \sum_y xy \, P(X = x, Y = y)$

Continuous RVs

$P(a \le x \le b) = \int_a^b f(x)\,dx$ → pdf



$P(x = 0) \to P(a \le x \le b) = P(a < x < b)$
$= P(a \le x < b) = P(a < x \le b)$

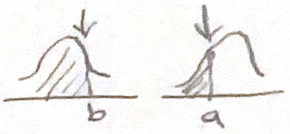
Cumulative Distribution Function

$F(x) = P(X \le x) = \int_{-\infty}^{x} f(y)\,dy$

$P(X > a) = 1 - F(a)$

$P(a \le x \le b) = F(b) - F(a)$

PDF $\underset{\text{integrate}}{\overset{\text{derive}}{\leftrightarrows}}$ CDF



Expected Value

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

$E[h(x)] = \mu_{h(x)} = \int_{-\infty}^{\infty} h(x) f(x)\,dx$

Variance

$V(X) = E[X^2] - E[X]^2$

Joint Distribution

$P(x, y) = \iint f(x, y)\,dx\,dy$

$P(a \le x \le b, c \le y \le d) = \int_a^b \int_c^d f(x, y)\,dy\,dx$

$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\,dy$ ↳ integrate over y

↓ marginal of x

$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,dx$

Condition for Independence

$f_{X,Y}(x, y) = f_X(x) f_Y(y)$

$E[g(x, y)] = \iint \underbrace{g(x, y)}_{x \cdot y} f(x, y)\,dx\,dy$

$E[xy] \longrightarrow x \cdot y.$

Expected Values, Covariance, Correlation

if independent, corr = 0
if corr = 0, not necessarily independent

Expected Values

Discrete RVs

- $E[x] = \sum x P_x(x)$
- $E[g(x)] = \sum g(x) p_x(x)$
- $E[n(x,y)] = \sum n(x,y) P_{x,y}(x,y)$
  ↳ joint pmf

Continuous RVs

- $E[x] = \int_{-\infty}^{\infty} x f_x(x)\, dx$
- $E[g(x)] = \int_{-\infty}^{\infty} g(x) f_x(x)\, dx$
- $E[n(x,y)] = \iint n(x,y) f_{x,y}(x,y)\, dx\, dy$
  ↳ joint pdf

Covariance → how strongly two RVs change

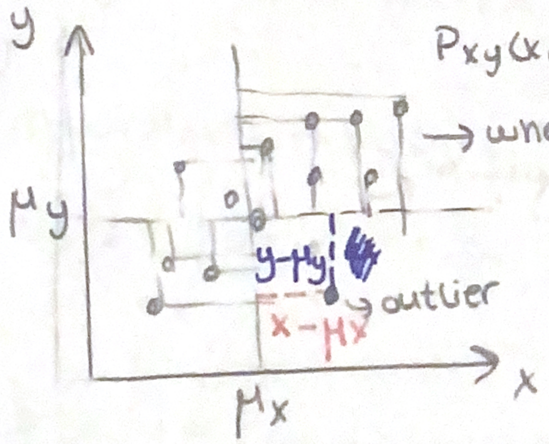
$E[x] = \mu_x$ X & Y are joint random variables.

$E[y] = \mu_y$

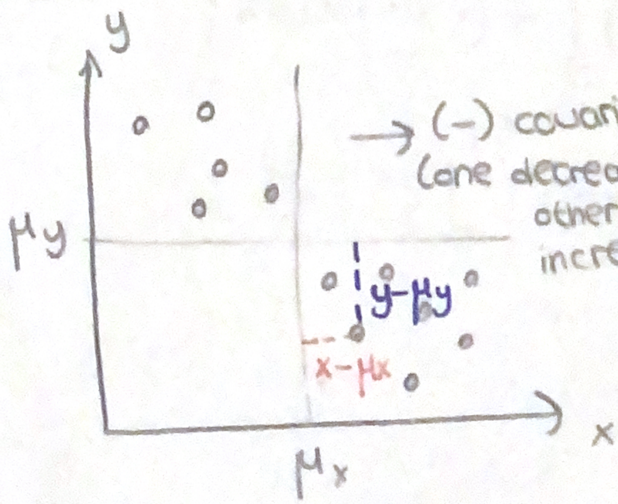
$$\mathrm{Cov} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{N-1}$$

$$\mathrm{Cov}(x,y) = E[(x-\mu_x)\cdot(y-\mu_y)]$$

- For Discrete X & Y → $\sum_x \sum_y (x-\mu_x)(y-\mu_y) P_{xy}(x,y)$
- For Continuous X & Y → $\iint (x-\mu_x)(y-\mu_y) f_{xy}(x,y)\, dx\, dy$



→ when x is (+) and (−) → multiplication gives same sign, everything is positive
y is (+) (−)



→ (−) covariance (one decreases, other one increases)

Covariance formula has units (e.g. kg × m) we need a metric for magnitude.

- $\rho_{x,y} = \frac{\mathrm{Cov}(x,y)}{\sigma_x \sigma_y}$ } correlation

$$= \frac{E[(x-\mu_x)(y-\mu_y)]}{\sqrt{E[(x-\mu_x)^2]\, E[(y-\mu_y)^2]}}$$

$\rho_{x,x} = 1$
$\rho_{x,-x} = -1$ } perfectly correlated
↳ $y = -x$

What if 2 vars are independent?
$\rho_{x,y} = 0$ (But not the other way around)

Bayesian Statistics

Bayes Rule

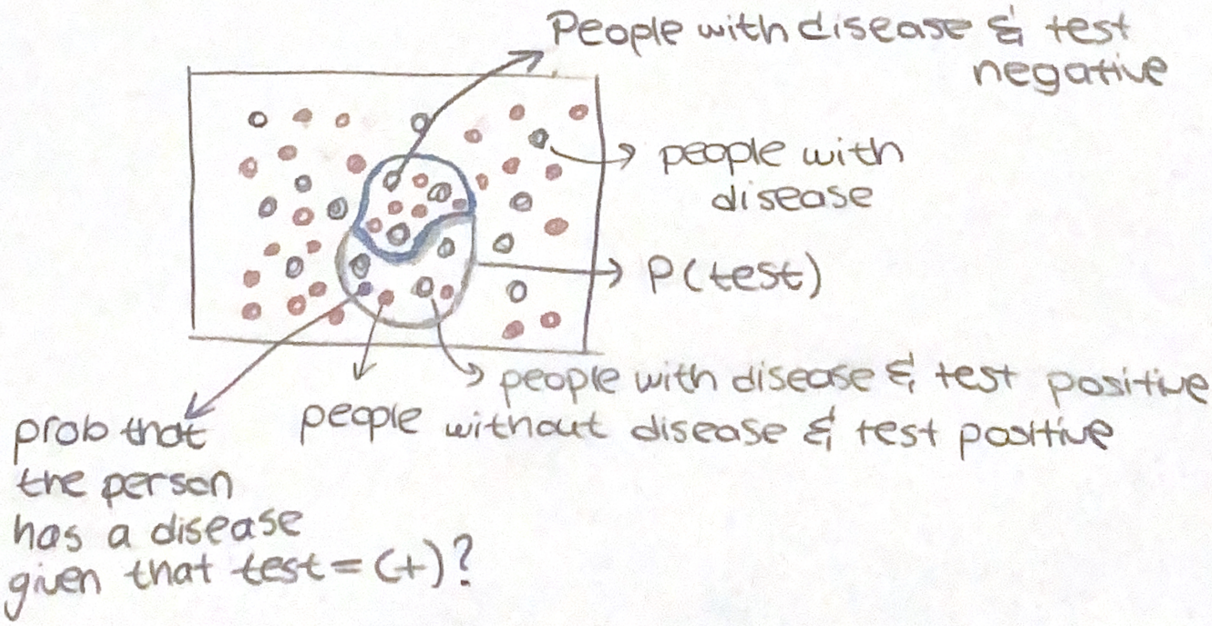
$$P(A \,\&\, B) = P(A|B)\,P(B) = \text{Likelihood} \times \text{Prior}$$

$$\underbrace{P(B|A)}_{\text{Posterior}} = \frac{P(A \,\&\, B)}{P(B)} = \frac{P(A|B)\,P(B)}{\underbrace{P(B)}_{\text{Evidence}}}$$

$$P(\text{disease}|\text{test}(+)) = \frac{P(\text{disease} \,\&\, \text{test}(+))}{P(\text{test}(+))}$$

→ people with disease and were tested (+)

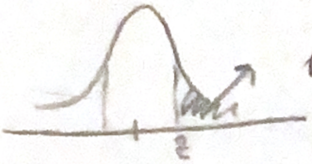
$$= \frac{P(\text{test}(+)|\text{disease})\,P(\text{disease})}{P(\text{test}(+))}$$



z test

1. Bulunacak parametreyi seç.
2. Ho, Ha & α'yı tanımla.
3. Test statistic'i bul (z ya da t)
4. $\bar{x}$, s, n bul, z'de yerine koy
5. P value'yu bul.
6. P value & α'yı karşılaştır.

Upper Tailed Test



Ha contains >

$1-\Phi(z)=P$

Lower Tailed Test

Ha has <

$P\text{-value}=\Phi(z)$

Two Tailed Test

$\neq$

$P\text{-value}=2[1-\Phi(|z|)]$

exe $\alpha=0.05$ $n=19$ $\bar{x}=562.68$

$H_o=\mu=500$ $\frac{s}{\sqrt{n}}=41.495$

$H_a:\mu>500$ $s=180.874$

$T=\frac{\bar{x}-500}{\frac{s}{\sqrt{n}}} \rightarrow t=\frac{562.68-500}{41.495}$

$=1.51$

$v=19-1=18$

t \ v	18
1.5	0.075

$1.5 \rightarrow 0.075 > 0.05 \rightarrow$ fail to reject Ho

Errors

fail to reject if $p>\alpha$

0 p 1

Type I error → Rejection Ho when it was true (α)

Type II error → Not rejecting Ho when it was false (β)

$$\chi^2=\sum\frac{(O-E)^2}{E}$$

O = Observed

E = expected

k = # kategori

$v=k-1$

z - test	t - test
n büyük	n küçük
σ biliniyor	σ bilinmiyor (?)
$\bar{x}$ normal	$\bar{x}$ normal
bilinmiyorsa s kullan	
$z=\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$	$t=\frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$

t_{n-1} curve

t \ v	(18) n-1
1.5	→ 0.075

p value

z	.06
1.9	0.9750

$\Phi(z)=P(Z\leq z)$

P-value > α → fail to reject Ho

exe $H_a:\mu>100$ $\alpha=0.05$

$H_o:\mu=100$, $\sigma_{\bar{x}}=2$ ↳ $\frac{\sigma}{\sqrt{n}}$

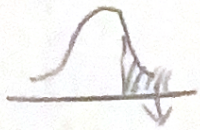
$\bar{x}=103$

$z=\frac{103-100}{2 \rightarrow \frac{\sigma}{\sqrt{n}}}=1.5$

z = Number of σs between μ_0 & $\bar{x}$

z ↑ Ho rejection ↑

P-value $=P(Z\geq z$ when Ho is true)



Ha: $\mu>100$

$1-\Phi(1.5)=0.0668$

P-value > α

$(0.0668)>0.05$

fail to reject

Confidence Interval

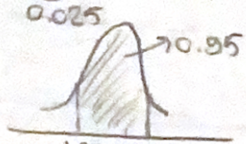
$\alpha=1-CL$ 0.95

$z_{\alpha/2}=1.96$

0.025

$\bar{x}-1.96\frac{\sigma}{\sqrt{n}}$ $\bar{x}$ $\bar{x}+1.96\frac{\sigma}{\sqrt{n}}$

$z_{\alpha/2}$



0.95

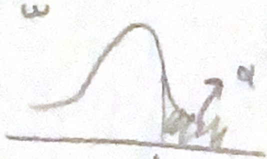
-1.96 1.96

Confidence Interval (small n) $\rightarrow P(-t_{\alpha/2,n-1}<T<t_{\alpha/2,n-1})=1-\alpha$

CI with t $\rightarrow \bar{x}\pm t_{\alpha/2,n-1}\cdot\frac{s}{\sqrt{n}}$

CI with z $\rightarrow \left(\bar{x}-1.96\frac{\sigma}{\sqrt{n}}, \bar{x}+1.96\frac{\sigma}{\sqrt{n}}\right)$

Sample size for width w $\rightarrow n=\left(2z_{\alpha/2}\frac{\sigma}{w}\right)^2$



α

$t_{\alpha,v}$

> t critical value

$p\leq\alpha \rightarrow$ reject

$p\geq\alpha \rightarrow$ fail to reject

exa (Hypothesis Testing)

Playing H or T with a friend.
I lost 55% of tosses.
$\alpha = 0.01$, $N > 40$
At what n do I need to decide that he is manipulating the coin?

H_0: He is not manipulating the coin toss.
H_a: He is manipulating the coin toss.

1. part

If p-value < %1 (0.01) → H_a is true

Bernoulli trial

X 1: win $1/2$
0: loss $1/2$

Ho says these are probs

$E[x] = 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} = \frac{1}{2}$ ($\mu \to \frac{1}{2}$)

$E[x^2] = 1^2 \frac{1}{2} + 0^2 \frac{1}{2} = \frac{1}{2}$

$\sigma^2 = E[x^2] - E[x]^2 = \frac{1}{4}$

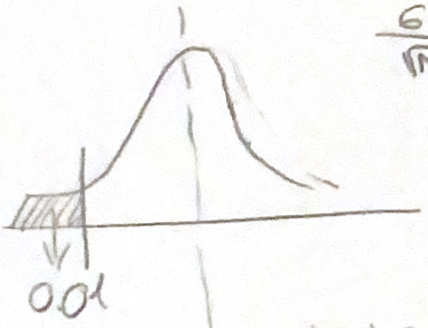
$\sigma = \frac{1}{2}$

$\mu = \frac{1}{2}$ $\sigma = \frac{1}{2}$ $\bar{x} = \frac{x_1 + x_2 + \dots x_N}{x_N}$ $\mu_{\bar{x}} = \mu$ $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}}$

$\bar{x} = 0.45$

Find z value $= \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{N}}} = \frac{0.45 - 0.5}{\frac{1/2}{\sqrt{N}}} = -\frac{\sqrt{N}}{10} = z$

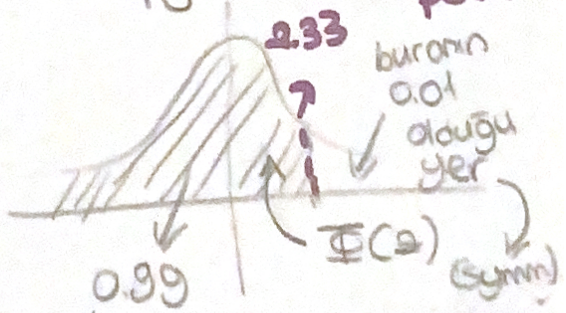
2nd part



→ Find critical z value where area under curve is 0.01

$z_c = -\frac{\sqrt{N}}{10}$

3rd part



z	0.00	0.01	0.02	0.03
0.0				
0.1				
⋮				
2.3				0.9901

↓ we know area, we lookup z

$z_c = -\frac{\sqrt{N}}{10}$ → left tail, use (+) for right

$2.33 = \frac{\sqrt{N}}{10}$

↓ when z is this H_a is true

$\underline{\underline{N = 543}}$