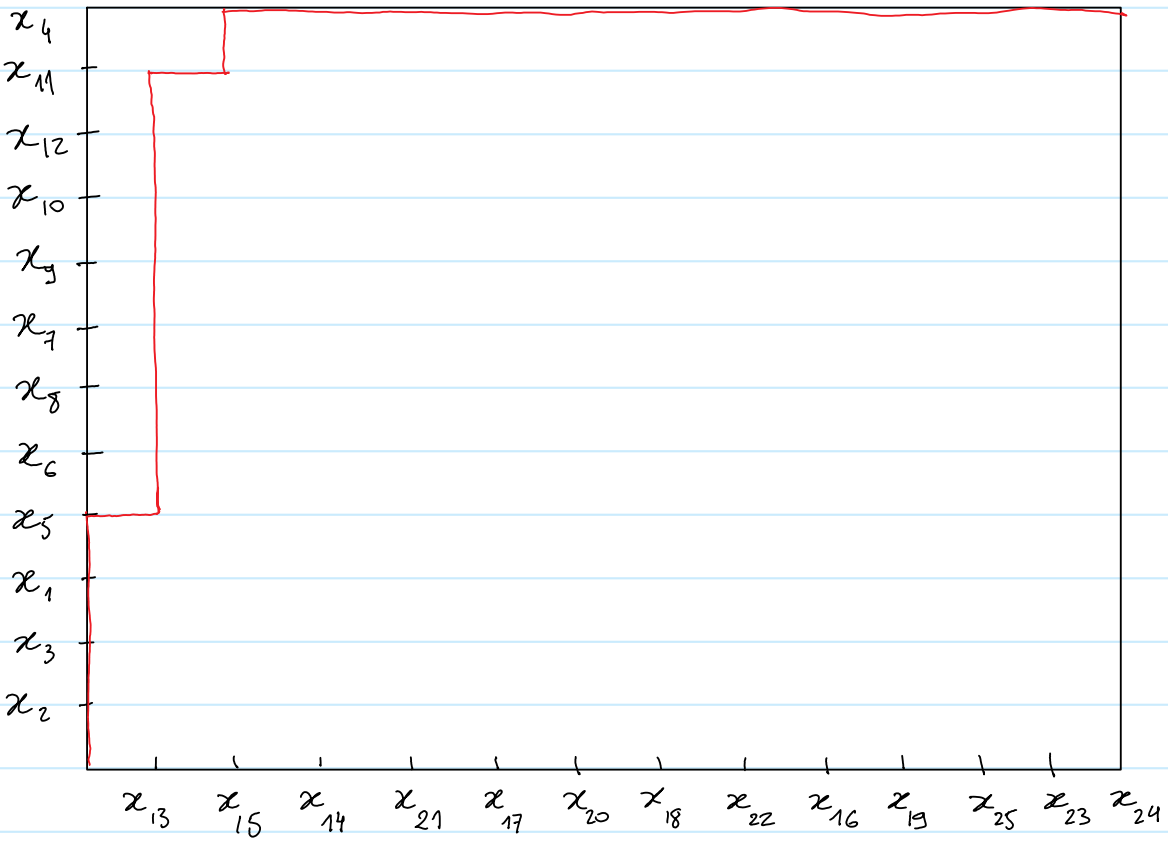
I a) $8 + 1 = 9$

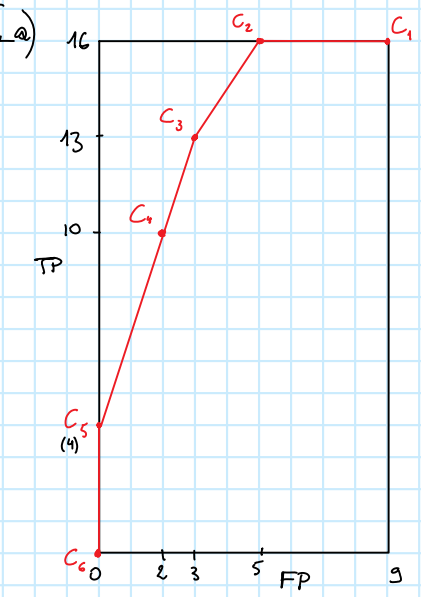
b) $\frac{9}{12 \cdot 13} \simeq 0.0577$

c) $1 - 0.0577 \simeq 0.9423$

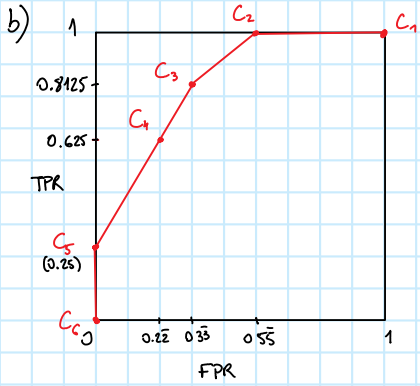
d)



II a)



b)



	C1	C2	C3	C4	C5	C6
TP	16	16	13	10	4	0
FP	9	5	3	2	0	0
Accuracy	0.64	0.8	0.76	0.68	0.52	0.36
Precision	0.64	0.761905	0.8125	0.833333	1	Undefined
Recall	1	1	0.8125	0.625	0.25	0

c) Max accuracy: C_2, min accuracy: C_6

d) Max precision: C_5, min precision: C_1 (or C_6 undefined).

e) Max recall: C_1, C_2, min recall: C_6

f) C_2, C_1

g) C_5, C_6

III $d = \frac{|3x_1 + 2x_2 + 4x_3 - 18|}{\sqrt{29}}$. Margin is computed as $z(x) = c(x_i) \cdot \frac{(3x_1 + 2x_2 + 4x_3 - 18)}{\sqrt{29}}$

Spreadsheet

1 if $z \leq 0$, 0 else

$\max(0, (1-z))$

$(1-z)^2$ if $z < 1$, 0 else

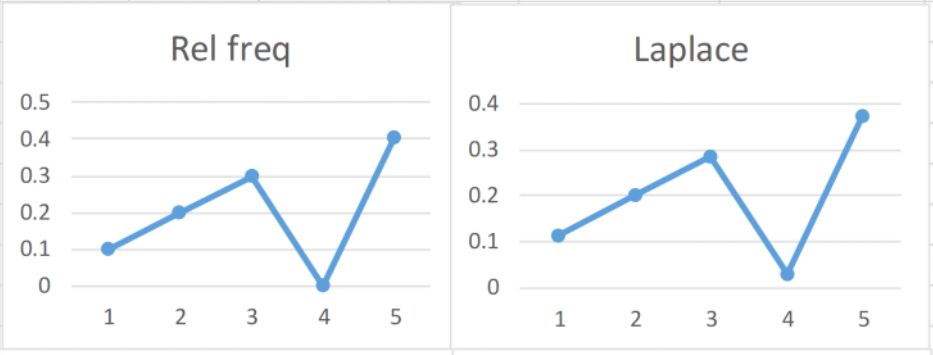
x1	x2	x3	c(x)	margin	0-1 loss	hinge	squared
2	2	3	1	0.742781	0	0.257219	0.066161
3	3	2	1	0.928477	0	0.071523	0.005116
1	2	3	1	0.185695	0	0.814305	0.663092
1	4	1	1	-0.55709	1	1.557086	2.424517
4	4	4	1	3.342516	0	0	0
2	2	2	1	0	1	1	1
3	3	1	-1	-0.1857	1	1.185695	1.405873
1	1	1	-1	1.671258	0	0	0
3	3	2	-1	-0.92848	1	1.928477	3.719022
0	4	2	-1	0.371391	0	0.628609	0.39515
4	0	0	-1	1.114172	0	0	0
0	0	3	-1	1.114172	0	0	0

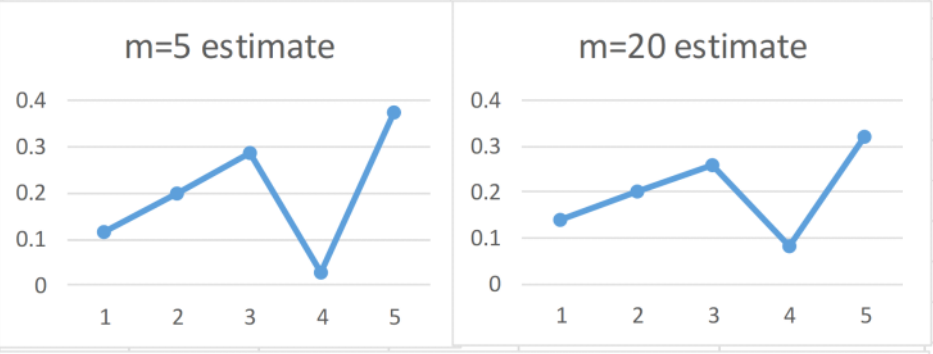
IV

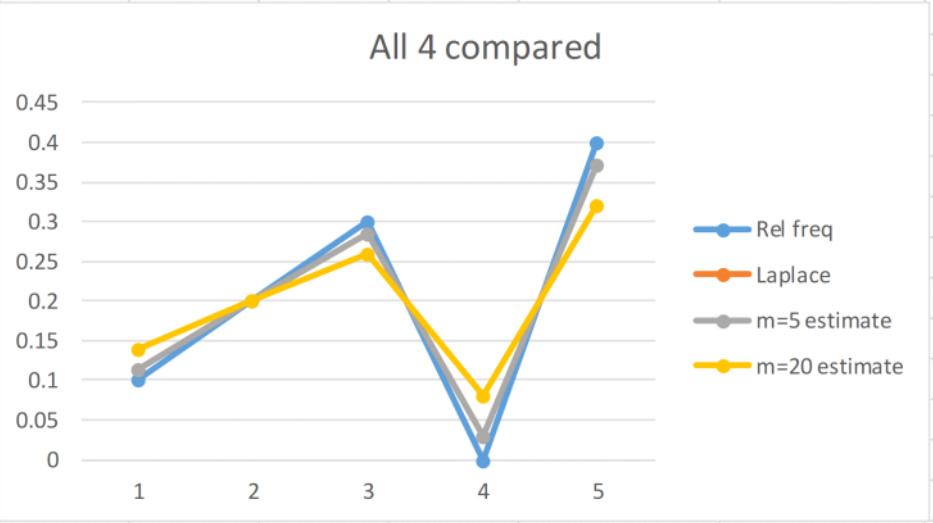
Increasing pseudo-count flattens curves (makes probs uniform: $\lim_{m\to\infty} \frac{N_i + m\cdot\pi_i}{|S|+m} = \pi_i$ and here $\pi_i = 1/k$ so makes all probs equal to 1/5)

Spreadsheet

Class	Count	Rel freq	Laplace	m=5 estimate	m=20 estimate
1	3	0.1	0.114286	0.114285714	0.14
2	6	0.2	0.2	0.2	0.2
3	9	0.3	0.285714	0.285714286	0.26
4	0	0	0.028571	0.028571429	0.08
5	12	0.4	0.371429	0.371428571	0.32







I a) The cardinality of the features set is $|F| = 3 \cdot 2 \cdot 2 \cdot 3 \cdot 4 = 144$
Therefore there are 2^{144} hypothesis to test $\Rightarrow t = 2^{144} \cdot 10^{-9}\,s = 2.23 \cdot 10^{34}\,s.$

b) Add X to each feature. $t = 4 \cdot 3 \cdot 3 \cdot 4 \cdot 5 \cdot 10^{-9}\,s = 7.2 \cdot 10^{-7}\,s.$

c) $t = (2^3-1)(2^2-1)(2^2-1)(2^3-1)(2^4-1) \cdot 10^{-9}\,s$

$t = 6.615 \cdot 10^{-6}\,s.$