Devon Callanan Pattern Recognition Professor Can

1. a) binomial distribution : $f(k,n,p) = \binom{n}{k} p^k (1-p)^{n-k}$

$k=0 \quad n=10$

	$f(k,n,p)$ $m=1$	$1-(1-f(k,n,p))^m$ $m=1{,}000$	$m=1{,}000{,}000$
$p=.04$	$.665$	$1-(1-.665)^{1000} \approx 1$	≈ 1
$p=.075$	$.459$	$1-(1-.459)^{1000} \approx 1$	≈ 1

b) Attached

2

$$f^*(x) = \arg\max_k \mu_k(x) = \arg_k\max \pi_k g_k(x)$$

having decision regions such that $\Gamma_k(f) := \{x : f(x) = k\}$

we want to minimize Risk, or maximize success so $1-R(f) = \mathbb{P}[f(x) = Y]$

$$1-R(f) = \sum_{k=0}^{K-1} \pi_k \times \mathbb{P}[f(x) = k \mid Y = k]$$

$$= \sum_{k=0}^{K-1} \pi_k \times \sum_{x: f(x)=k} g_k(x) = \sum_{k=0}^{K-1} \pi_k \times \sum_{\Gamma_k}$$

we must pick regions Γ_k such that the prob $(\pi_k g_k(x))$ is maximized

so $f^*(x) = \arg\max_k \pi_k g_k(x)$

3 a) Attached

b) $\pi_0 = \pi_1$ so $f^*(x) = \arg_k\max g_k(x)$ $\therefore$ $\frac{1}{\sqrt{2}} e^{-\sqrt{2}|x|} \overset{0}{\gtrless} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

Bayes Rule: $\sqrt{\pi}\, e^{-\sqrt{2}|x| + x^2/2} \overset{0}{\gtrless} 1$

*ignoring outer bound. Risk assumed negligiable * $\pi_0 = \pi_1 = \frac{1}{2}$?

c) $$R(f_*) = \pi \int_{-\infty}^{.495} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx + \pi \int_{.495}^{\infty} \frac{1}{\sqrt{2}} e^{-\sqrt{2}|x|} dx = \frac{1}{2}(.6897) + \frac{1}{2}(0.248) = \boxed{0.46885}$$

4) $a = \hat{\Sigma}^{-1}(\hat{\mu}_0 - \hat{\mu}_1)$ $\qquad b = -\frac{1}{2}\hat{\mu}_0^T\hat{\Sigma}^{-1}\hat{\mu}_0 + \frac{1}{2}\hat{\mu}_1^T\hat{\Sigma}^{-1}\hat{\mu}_1 + \log\frac{\hat{\pi}_0}{\hat{\pi}_1}$

$$f(x) = \arg\max_k \hat{\pi}_k \, \phi(x : \hat{\mu}_k, \hat{\Sigma})$$

log

$$= \arg\max_k \log\hat{\pi}_k - \frac{1}{2}(x - \hat{\mu}_k)^T\hat{\Sigma}^{-1}(x - \hat{\mu}_k)$$

distribute

$$= \arg\max_k \log\hat{\pi}_k - \frac{1}{2}\hat{\mu}_k^T\hat{\Sigma}^{-1}\hat{\mu}_k - \frac{1}{2}x^T\Sigma^{-1}x + x^T\hat{\Sigma}^{-1}\mu_k$$

in case $K=2$, decision rule becomes

$$\log\hat{\pi}_0 - \frac{1}{2}\hat{\mu}_0^T\hat{\Sigma}^{-1}\hat{\mu}_0 - \cancel{\frac{1}{2}x^T\hat{\Sigma}^{-1}x} + x^T\hat{\Sigma}^{-1}\mu_0 \underset{1}{\overset{0}{\gtrless}} \log\hat{\pi}_1 - \frac{1}{2}\hat{\mu}_1^T\hat{\Sigma}^{-1}\mu_1 - \cancel{\frac{1}{2}x^T\hat{\Sigma}^{-1}x} + x^T\hat{\Sigma}^{-1}\mu_1$$

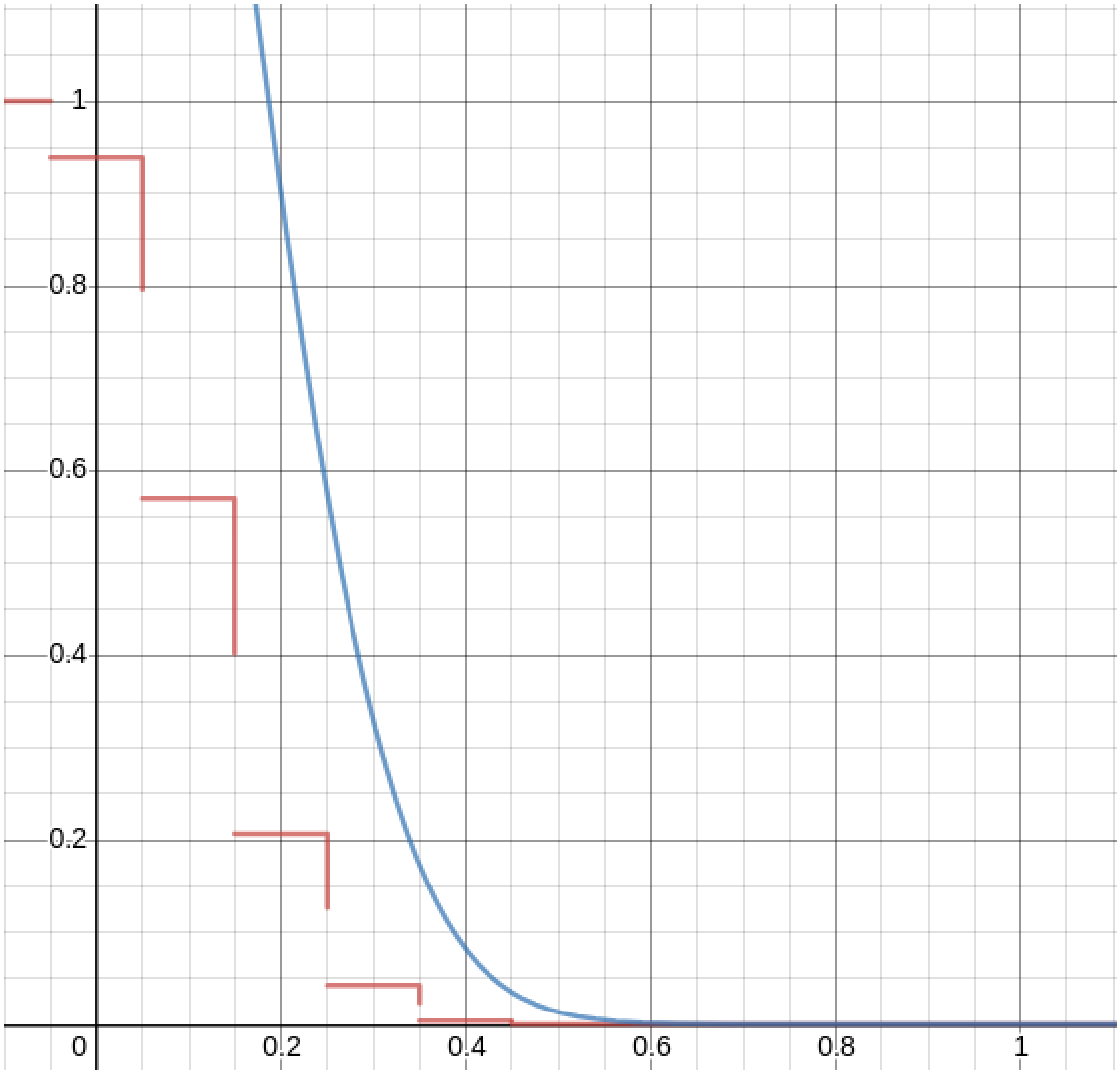
simplify

$$\log\hat{\pi}_0 - \log\hat{\pi}_1 - \frac{1}{2}\mu_0^T\Sigma^{-1}\mu_0 + \frac{1}{2}\mu_1^T\Sigma^{-1}\mu_1 + x^T\Sigma^{-1}\mu_0 - x^T\Sigma^{-1}\mu_1 \underset{1}{\overset{0}{\gtrless}} 0$$

$$\underbrace{\log\frac{\hat{\pi}_0}{\hat{\pi}_1} - \frac{1}{2}\hat{\mu}_0^T\hat{\Sigma}^{-1}\hat{\mu}_0 + \frac{1}{2}\hat{\mu}_1^T\hat{\Sigma}^{-1}\hat{\mu}_1}_{b} + x^T\underbrace{\hat{\Sigma}^{-1}(\mu_0 - \mu_1)}_{a} \underset{1}{\overset{0}{\gtrless}} 0$$

$$x^T\Sigma^{-1}(\mu_0 - \mu_1) = [\Sigma^{-1}(\mu_0 - \mu_1)]^T x$$

1. b)

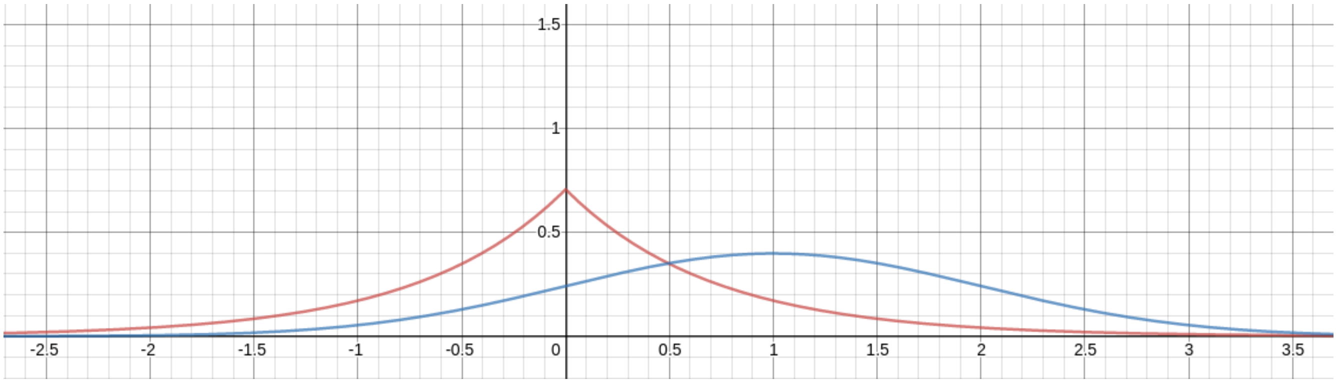


1 $$1-\left(\sum_{n=5-10x}^{5+10x} \mathrm{nCr}(10,n)\cdot .5^{10}\right)^2$$

2 $$2\cdot e^{-\left(2x^2\cdot N\right)}$$

3 $$N=10$$

3. a)



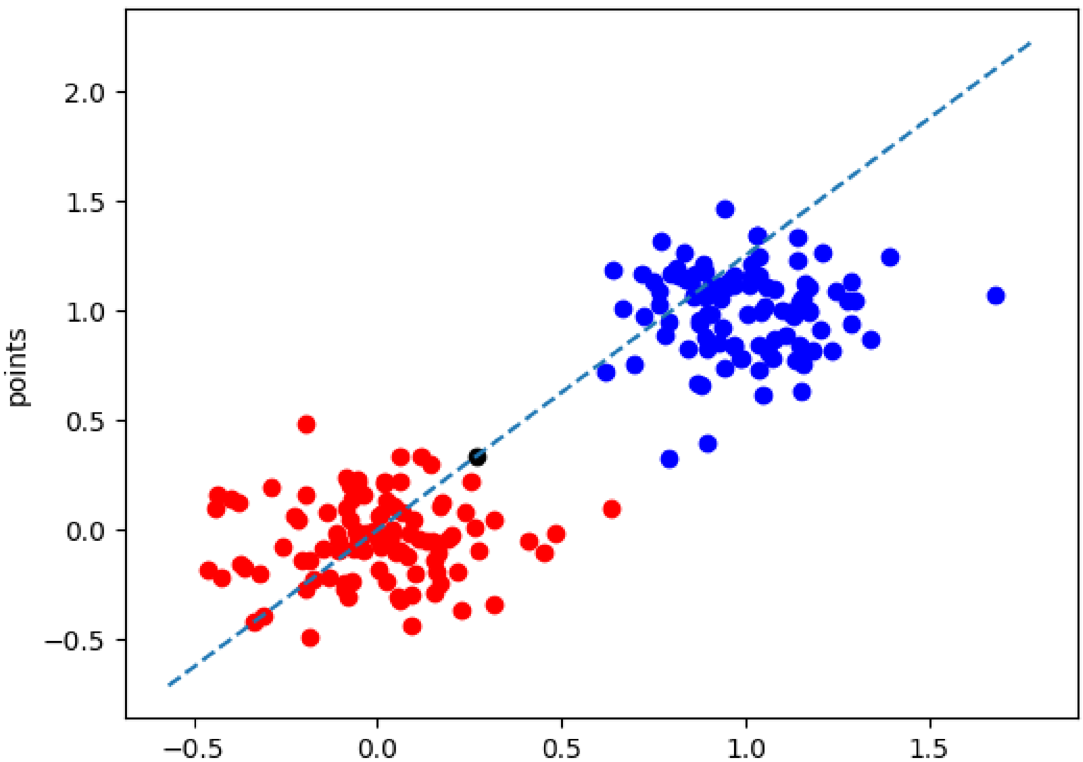
5.

	Synthetic 1	Synthetic 2	Synthetic 3	Synthetic 4	Real
Normal	0.995	0.99	0.78	0.75	0.7
Modified	1	0.99	0.7	0.75	0.67

The modified co-variance strategies has a beneficial effect on some, but not all, of the data sets. The performance of LDA is strongly coupled with the separability of the data. Groupings with tight co-variance and strong inter group variance perform best. The diagrams bellow help shed light on the performance statistics shown above.

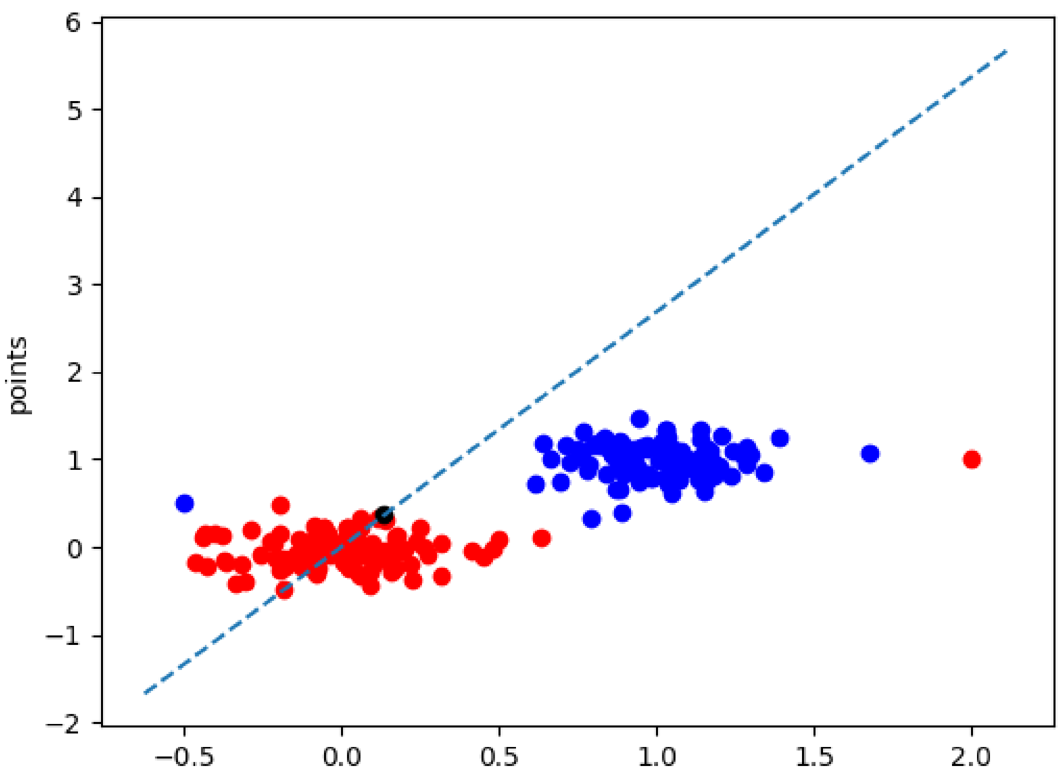
Plots of data sets and decision rules:

1.



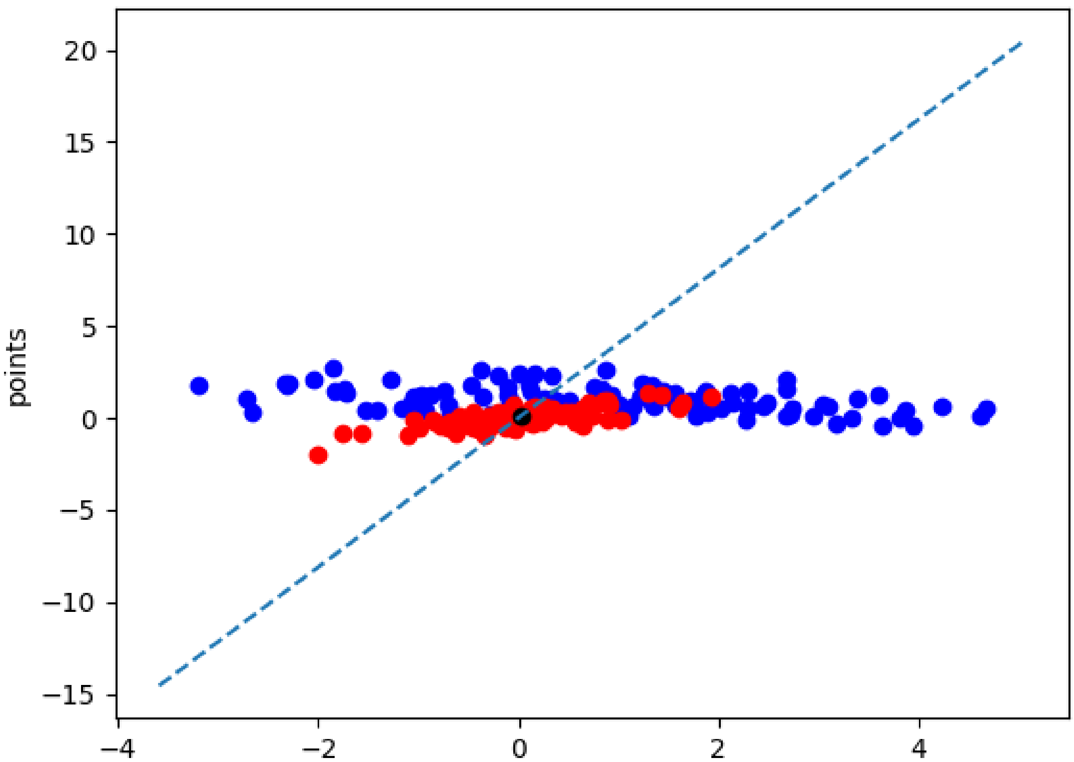
points
2.0
1.5
1.0
0.5
0.0
−0.5
−0.5
0.0
0.5
1.0
1.5

2.



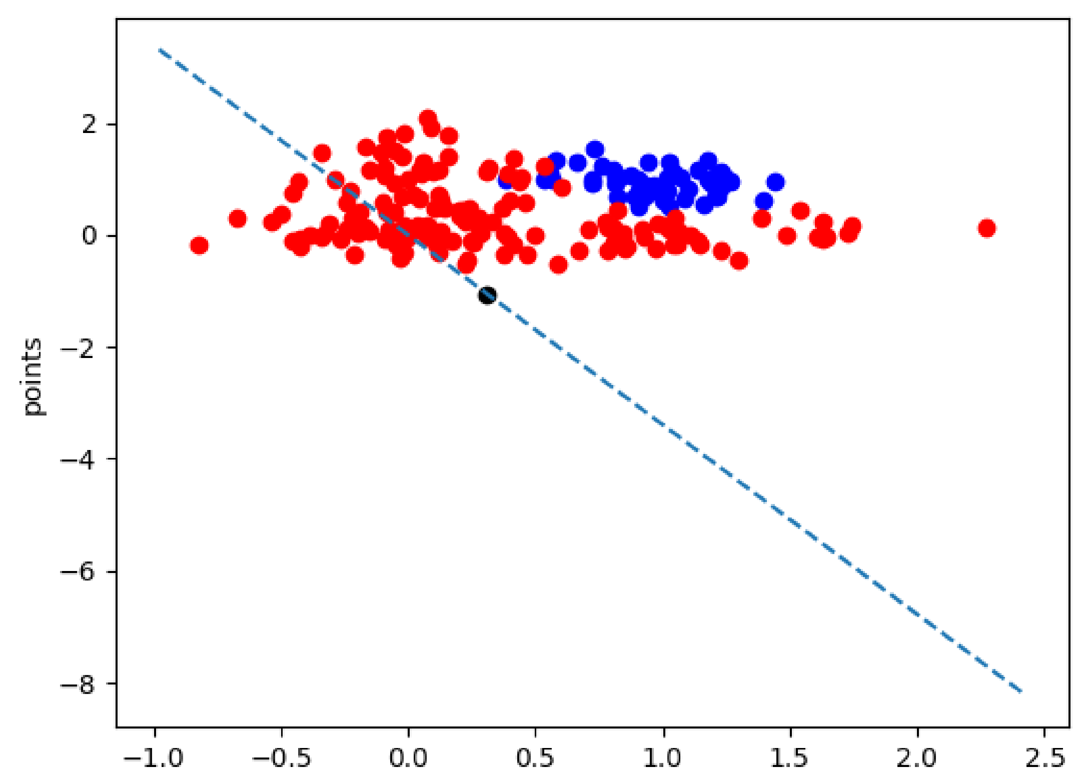
points
6
5
4
3
2
1
0
−1
−2
−0.5
0.0
0.5
1.0
1.5
2.0

3.



points
20
15
10
5
0
−5
−10
−15
−4
−2
0
2
4

4.



points
2
0
−2
−4
−6
−8
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
2.5