LaTeX Mini Project 2

Sachin Srivastava
Rutgers Univeristy
sachin.srivastava@rutgers.edu

1. Problem 1

The complete strip is at most ϵ.
Probability that we miss a strip is $1 - \epsilon$.
Probability that N instances miss a strip is $(1 - \epsilon)^n$.

We want, $(1 - \epsilon)^n <= \delta$

Now, we know that $(1 - x) <= \exp^x$.
Therefore, $\exp^{-N*\epsilon} <= \delta$

Which implies, $N >= \frac{1}{\epsilon} Ln \frac{1}{\delta}$

2. Problem 2

2.1.

For a finite and consistent hypothesis $\mathbb{E}[h_d] <= \frac{1}{N}(Ln|H| + Ln\frac{1}{\delta})$

Here,
H=2800
N=200
$1 - \delta = 0.95$
$\delta = 0.05$

Which implies
$\epsilon <= 1/200(Ln|2800| + Ln(1/0.05))$ Therefore,
$\epsilon <= 0.054$

2.2.

Now since the hypothesis is inconsistent we have $\mathbb{E}[h_d] <= \mathbb{E}^{'}[h|d] + \sqrt{\frac{Ln|H| + Ln(2/\delta)}{2N}}$
We have,
$\epsilon <= 20/200 + \sqrt{\frac{Ln2800 + Ln(2/0.05)}{2*200}}$
Therefore,
$\epsilon <= 0.27$

3. Problem 3

3.1.

$P(error) = \int_{-\infty}^{\infty} P(error, x)dx$
$= \int_{-\infty}^{\infty} P(error|x)P(x)dx$

Thus,
$P(error) = P(w_1 \; and \; x < \theta) + P(w_2 \; and \; x > \theta)$
$= P(x < \theta|w_1)P(w_1) + P(x > \theta|w_2)P(w_2)$

Therefore,
$P(error) = \int_{-\infty}^{\theta} P(x|w_1)P(w_1)dx + \int_{\theta}^{\infty} P(x|w_2)P(w_2)dx$

3.2.

Differentiating the above equation,
$dP(error)/dx = P(w_1)[p(x|w_1)]_{-\infty}^{\theta} + P(w_2)[p(x|w_2)]_{\theta}^{\infty}$
$= P(w_1)P(\theta|w_1) - P(w_2)P(\theta|w_2)$
As at ∞ and $-\infty$, $P(x|w) = 0$

Equating $dP(error)/dx$ to zero to find the minima we get,
$P(w_1)P(\theta|w_1) = P(w_2)P(\theta|w_2)$

3.3.

Here, θ is not necessarily a unique value. This inequality might be true for different values of x and also θ can be a range of values.

3.4.

If we take two symmetric Gaussian distribution with equal variance but different means. At there point of intersection the P(error) will be maximum. For example if we take N(0,1) and (2,1) then at θ equal to 1, the P(error) will be maximum.

4. Problem 4

4.1.

For a deterministic decision rule we have overall risk $R = \int R(\alpha|x)p(x)dx$
For a stochastic decision rule we have $P(\alpha_i|x)$ of deciding each action $\alpha_i(x)$
Then,
$R(\alpha(x)|x) = \sum_{i=1}^{a} R(\alpha_i(x)|x)P(\alpha_i|x)$
Therefore, the overall risk in this case is

$$R(\alpha) = \int [\sum_{i=1}^{a} R(\alpha_i(x)|x)P(\alpha_i|x)]p(x)dx$$

4.2.

If we have, $a_1 + a_2 + ... + a_n = 1$ and $a_1, a_2, ..., a_n > 0$
then, $a_1x_1 + a_2x_2 + ... + a_nx_n > min(x_1, x_2..., x_n)$
Using this we can say in
$\sum_{i=1}^{a} R(\alpha_i(x)|x)P(\alpha_i|x) > min[R(\alpha_i(x)|x)]_{i=1}^{a}$
Hence R is minimized by choosing the $min[R(\alpha_i(x)|x)]$ with $P(\alpha_i|x) = 1$

4.3.

Yes, one can gain from randomizing a sub optimal rule because if a lower risk action happens with high probability, the overall risk will lower.

5. Problem 5

5.1.

Let $a <= b$ then, $b = a + \delta$, where $\delta >= 0$
Therefore,
$\sqrt{ab} = \sqrt{a(a+\delta)} = \sqrt{a^2 + a\delta} > a$
$\therefore \sqrt{ab} >= a$
Therefore we have,
$min[a, b] <= \sqrt{ab}$

5.2.

$P(error) = \int min[P(w_1)P(x|w_1), P(w_2)P(x|w_2)]dx$

From the proof in the above we get,
$P(error) <= \int \sqrt{P(w_1)P(w_2)}\sqrt{P(x|w_1)P(x|w_2)}]dx$
Therefore,
$P(error) <= \sqrt{(w_1)P(w_2)} \int \sqrt{P(x|w_1)P(x|w_2)}]dx$

Now, $\sqrt{P(w_1)P(w_2)} <= 0.5$
Hence, $P(error) <= 1/2\rho$, where,
$\rho = \int \sqrt{P(x|w_1)P(x|w_2)}dx$

5.3.

The Bhattacharyya error bound for two multivariate Gaussian densities with the unit covariance and uniform priors is 0.5.

This is because the $\int \sqrt{P(x|w_1)P(x|w_2)}dx = 1$ in this case.
Hence, $P(error) <= 1/2$