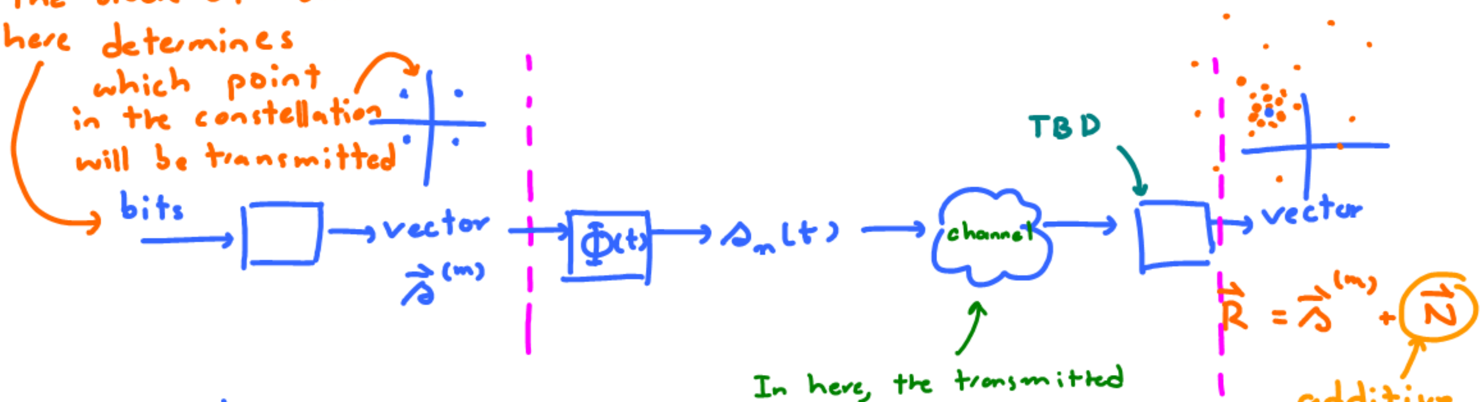
Vector Channel : Connection to waveform channel

The block of bits here determines which point in the constellation will be transmitted



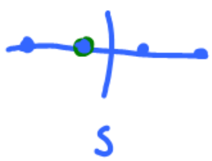
In here, the transmitted signal $s_m(t)$ is corrupted by noise $N(t)$.

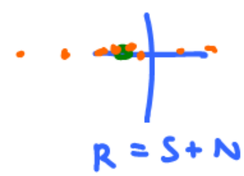
We will later revisit this "waveform model" later.

For now, we simply assume that the additive noise $N(t)$ here would produce additive noise vector $\vec{N}$ in the next step.

The connection between $\vec{N}$ and $N(t)$ requires knowledge on random process which is not our current focus.

1-D channel:





In chapter 6, we will consider 1-D channel: $R = S + N$

In chapter 7, we will extend the results from chapter 6 to the vector channel: $\vec{R} = \vec{S} + \vec{N}$.

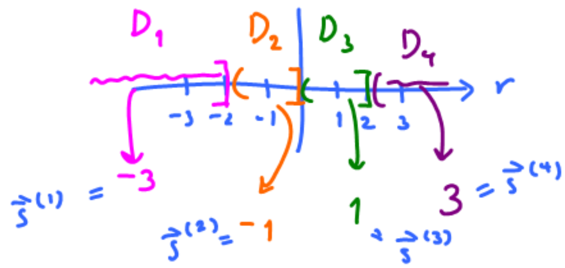
Detector $\hat{s}(\cdot)$

- input: output (of the the channel) $= R = S + N$.

  output: $\hat{S}$ = estimated value of the transmitted message.

- We consider deterministic detector. It can be shown that random detector does not help.

Ex. $\hat{s}(r) = \begin{cases} -3 & \text{if } r \le 2 \\ -1 & \text{if } -2 < r \le 0 \\ 1 & \text{if } 0 < r \le 2 \\ 3 & \text{if } r > 2 \end{cases}$



However, when you actually use the detector, its input (R) will be random. Therefore, its output will be random as well.

So, when we talk about the output of the detector, we write

$$\hat{S}$$

or $\hat{s}(R)$

They refer to the same thing.
The first one simply look at the estimate as a random variable.

The second one look at the estimate as the output of a deterministic detector when the input is random.

- Detection regions:

$$D_m = \text{set of } r \text{ whose } \hat{s}(r) = \vec{s}^{(m)}$$

Ex. In the example above,

D_2 is the interval (set of r) which, when recieved, the decoder $\hat{s}$ will give $s^{(2)}$ (which is -1).

6 Optimal Detection (Receiver) for Additive Noise Channels: 1-D Case

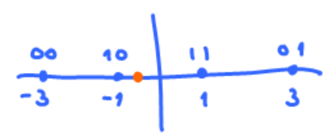
Definition 6.1. Detection Problem: Consider the problem of **detecting** the scalar message S in the presence of additive noise N. The received signal R is given by

$$R = S + N.$$

A **detector** uses R to predict the value of S. The predicted value is called $\hat{S}$. Because the detector works on R, it is a function of R and hence we may write the detector as $\hat{s}(\cdot)$ and write its output, which is the detected value, as $\hat{s}(R)$.

Detector is a function of R. Ex. $\hat{s}(r) = \begin{cases} -3 & \text{if } r \leq 2 \\ -1 & \text{if } -2 < r \leq 0 \\ 1 & \text{if } 0 < r \leq 2 \\ 3 & \text{if } r > 2 \end{cases}$

Definition 6.2. Probability of Error ($\mathcal{E}$): The goal of a detector is to produce $\hat{s}(R)$ that is the same as S. However, due to corruption by the noise N, this is not always possible. To measure the performance of a detector, we consider its probability of error: $p[\hat{S} \neq S]$



Performance Index: $$P(\mathcal{E}) = P[\hat{s}(R) \neq S].$$

Our goal is to theoretically predict the probability of error for a given detector $\hat{s}(\cdot)$. Because s is a symbol, the probability of error is also referred to as the **symbol error probability**.

Definition 6.3. Another type of error probability is the **bit error probability**. This error probability is denoted by P_b and is the error probability in transmission of a single bit.

Determining the bit error probability in general requires detailed knowledge of how different bit sequences are mapped to signal points. Therefore, in general finding the bit error probability is not easy unless the constellation exhibits certain symmetry properties to make the derivation of the bit error probability easy.

6.4. Gaussian Noise: We assume that the noise N is Gaussian with mean 0 and standard deviation σ_N. This implies that

$$f_N(n) = \frac{1}{\sqrt{2\pi}\sigma_N} e^{-\frac{1}{2}\left(\frac{n}{\sigma_N}\right)^2}.$$

$P[1 \leq N \leq 2] = \int_1^2 f_N(n)\,dn$

Definition 6.5. In general, a **Gaussian (normal) random variable** X with mean m and standard deviation σ is characterized by its probability density function (PDF):

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}. \quad \xrightarrow[\sigma=1]{m=0} \quad f_S(s) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}s^2}$$

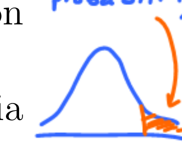
To talk about such X, we usually write $X \sim \mathcal{N}(m, \sigma^2)$. Probability involving X can be evaluated by

$$P[X \in A] = \int_A f_X(x)dx.$$

In particular,

$$P[X \in [a,b]] = \int_a^b f_X(x)dx = F_X(b) - F_X(a)$$

Usually, in communicating we deal with the "tail" probability.



where $F_X(x) = \int_{-\infty}^x f_X(t)dt$ is called the cumulative distribution function (CDF) of X.

We usually express probability involving Gaussian random variable via the Q function which is defined by

$$Q(z) = \int_z^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

Note that $Q(z)$ is the same as $P[S > z]$ where $S \sim \mathcal{N}(0,1)$; that is $Q(z)$ is the probability of the "tail" of $\mathcal{N}(0,1)$.

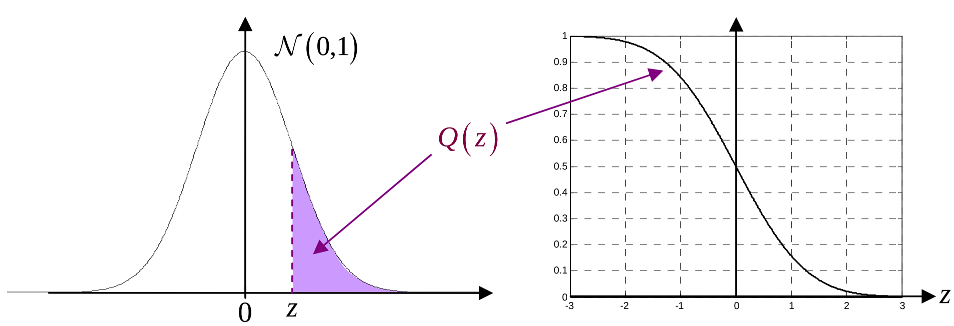


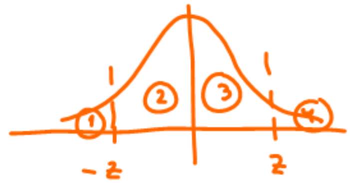
Figure 8: Q-function

It can be shown that

- Q is a decreasing function
- $Q(0) = \frac{1}{2}$

② + ③ + ④ = ① + ② + ③

- $Q(-z) = 1 - Q(z)$



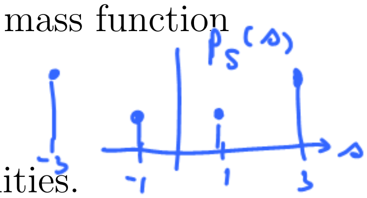
- For $X \sim \mathcal{N}(m, \sigma^2)$,

$$P[X > c] = Q\left(\frac{c-m}{\sigma}\right).$$

Ex. $\mathcal{S} = \{-3, -1, 1, 3\}$

6.6. Alphabet Set and Prior Probability: We will assume that S is (randomly) selected from an alphabet set $\mathcal{S}$ with probability mass function (PMF)

$$p_s = p_S(s) = P[S = s] \quad \text{for } s \in \mathcal{S}.$$



These probabilities are often referred to as the prior probabilities.

Example 6.7. For bipolar 1-D transmission, $\mathcal{S} = \{-s_0, s_0\}$ for some positive s_0.

Talking about optimality requires that you specify the objective function and whether you want to maximize or minimize. Here, we want to minimize P(E).

6.8. Optimal Detector: The optimal detection rule is the one that upon observing $R = r$ decides in favor of the message s that maximizes $P[S = s|R = r]$. In other words,

(See the next page for an informal proof of how the MAP detector is optimal.)

$$\hat{s}(r) = \arg\max_{s \in \mathcal{S}} P[S = s|R = r]. \tag{9}$$

form 1 : Use maximum a posteriori probability directly.
MAP

Definition 6.9. The detector defined by (9) is known as the **maximum a posteriori probability (MAP)** detector. Note that the MAP detector can be simplified to

Recall:

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

$$\hat{s}(r) = \arg\max_{s} \frac{f_{R|S}(r|s)\, P[S=s]}{f_R(r)}$$

$$P[S=s|R=r] = \frac{P[R=r|S=s]\, P[S=s]}{P[R=r]}$$

Need to use pdf for continuous R.

$$\hat{s}_{\text{MAP}}(r) = \arg\max_{s \in \mathcal{S}} p_s f_{R|S}(r|s) \tag{10}$$

form 2

Optimal Detector (Receiver)

(Derivation of MAP detector.)

error event

Goal: To mimize $P(\mathcal{E}) = P[\hat{S} \neq S] = P[\hat{s}(R) \neq S]$

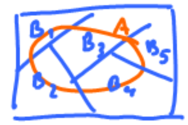
Recall: Total probability theorem

- nonoverlapping
- union of them gives Ω

Assume B_i's partition Ω

$$P(A) = \sum_i P(A \cap B_i)$$

$$= \sum_i P(A|B_i)P(B_i)$$



Consider events $B_r = [R = r]$

There are uncountably many values of r because R is a continuous R.V. So, the "summation" and $P(B_r)$ are not quite what we want to use here.

$$\left(P(\mathcal{E}) = \sum_r P[\hat{s}(R) \neq S \mid B_r] P(B_r) \right)$$

Need to use the pdf and $\int$ because R is continuous.

$$P(\mathcal{E}) = \int_r P[\hat{s}(R) \neq S \mid R = r] f_R(r)\, dr$$

Equivalently, minimizing $P(\mathcal{E})$ is the same as

Maximizing

$$\int_r P[\hat{s}(R) = S \mid R = r] \underbrace{f_R(r)}_{\text{nonnegative}}\, dr$$

Equivalent to

maximizing $P[\hat{s}(R) = S \mid R = r]$ for all r.

Equivalent to

$$\hat{s}(r) = \arg\max_s \underbrace{P[S = s \mid R = r]}_{\text{a posteriori probability}}$$

Form 1 of MAP Detector

The detector that does this is the optimal detector from the perspective of minimization of $P(\mathcal{E})$.

Recall: From ECS 315 that there are many names:

$P[S=s]$	$P[S=s \mid R=r]$
unconditional probability	conditional probability
prior probability	posterior probability
a priori probability	a posteriori probability

$(p_s = P[S = s]$ is the a priori probability.$)$

Because it maximizes a posteriori probability, we call it a MAP detector.

6.10. In the case where the messages are equiprobable a priori, the optimal detection rule in (10) reduces to

$$\hat{s}_{\mathrm{ML}}(r) = \arg\max_{s\in\mathcal{S}} f_{R|S}(r\,|s) \tag{11}$$

Definition 6.11. The term $f_{R|S}(r\,|s)$ in (11) is called the **likelihood** (or likelihood function) of message s, and the detector given by (11) is called the **maximum-likelihood detector** or **ML detector**.

- Note that the ML detector is not an optimal detector unless the messages are equiprobable.
- The ML detector, however, is a very popular detector since in many cases having exact information about message probabilities is difficult.

6.12. For additive noise channel where $R = S + N$ and $S \perp\!\!\!\perp N$,

$$P[R=r\,|\,S=s] = P[S+N=r\,|\,S=s] = P[s+N=r\,|\,S=s]$$
$$= P[N=r-s\,|\,S=s] = P[N=r-s]$$

Need to use pdf for continuous N.

$$f_{R|S}(r\,|s) = f_N(r-s)\,.$$

Therefore,

form 3

$$\hat{s}_{\mathrm{MAP}}(r) = \arg\max_{s\in\mathcal{S}} p_s f_N(r-s) \tag{12}$$

and

$$\hat{s}_{\mathrm{ML}}(r) = \arg\max_{s\in\mathcal{S}} f_N(r-s)\,. \tag{13}$$

6.13. For additive noise channel where $R = S + N$, $S \perp\!\!\!\perp N$, and $N \sim \mathcal{N}(0,\sigma^2)$

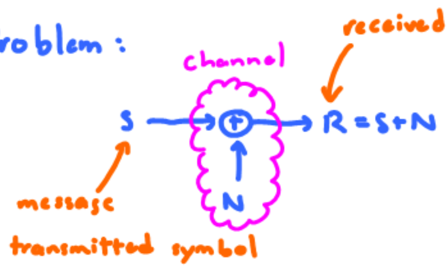
$$\hat{s}_{MAP}(r) = \arg\max_{s} p_s \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{r-s}{\sigma}\right)^2}$$

a constant (does not affect the maximization)

taking ln

$$\hat{s}_{\mathrm{MAP}}(r) = \arg\max_{s\in S}\left(2\sigma_N^2 \ln p_s - (r-s)^2\right)$$
$$= \arg\max_{s\in S}\left(\sigma_N^2 \ln p_s - \frac{\mathcal{E}_s}{2} + s\cdot r\right),$$

Recall: Problem:



At the receiver, $R = r$ is observed.

Estimate the value of S.

Assumption: S is randomly selected with pmf $p_s = P[S = s]$

$S \perp\!\!\!\perp N$

N is a continuous RV with pdf $f_N(n)$

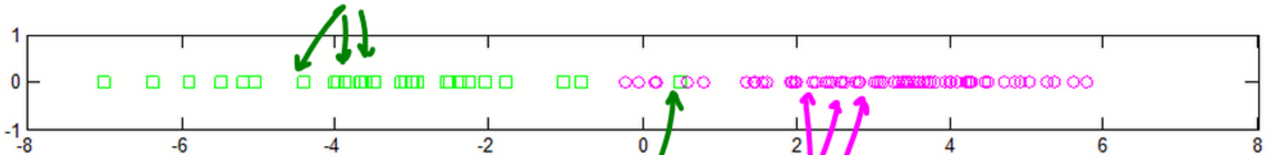
Example: The message S is randomly selected from the set $\{-3, 3\}$

with $p_3 = P[S=3] = 0.7$

$p_{-3} = P[S=-3] = 1 - 0.7 = 0.3.$

The message is corrupted by an independent additive N.

R when $S=-3$ was sent



R when $S=3$ was sent

with large noise, some R is very far away from $S=-3$.

Goal: Find a detector $\hat{s}(\cdot)$

$$R \to \boxed{\hat{s}(\cdot)} \to \hat{S}$$

to minimize $P(\mathcal{E}) = P[\hat{S} \neq S]$

error event

(same as maximizing $P[\hat{S} = S]$.)

Solution: The optimal detector is given by the MAP (maximum a posteriori) detector:

$$\hat{s}_{MAP}(r) = \arg\max_{s} \, p_s f_N(r - s)$$

When p_s is ignored, we have the maximum likelihood (ML) detector.

$$\hat{s}_{ML}(r) = \arg\max_{s} \, f_N(r - s)$$

Back to example: Only two possible s values $(3, -3)$

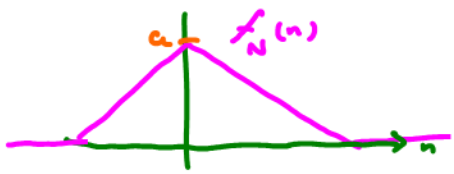
So, $\hat{s}(r) = \begin{cases} -3, & \text{when} \ldots \\ 3, & \text{when} \ldots \end{cases}$

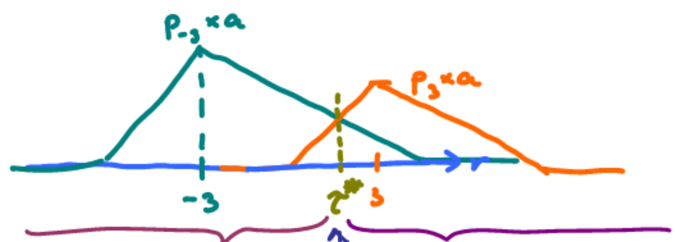
We should compare $p_s f_N(r-s)$

$$p_{-3} f_N(r-(-3)) \quad \text{vs.} \quad p_3 f_N(r-3)$$

$\hat{s}(r) = 3$ when $p_3 f_N(r-3) > p_{-3} f_N(r-(-3))$

Assume





on this side, $p_{-3} f_N(r-(-3)) > p_3 f_N(r-3)$. Therefore, $\hat{s}(r) = -3$

In this region, $p_3 f_N(r-3) > p_{-3} f_N(r-(-3))$. Therefore, $\hat{s}(r) = 3$

The case when $r = \tau^*$ occurs with zero probability. So, it does not matter which value the detector gives. For completeness, we usually include this case into one of the region.

$$\hat{s}_{MAP}(r) = \begin{cases} 3, & r > \tau^* \\ -3, & r \leq \tau^* \end{cases}$$

and

$$\hat{s}_{\mathrm{ML}}(r) = \arg\min_{s \in S} (r-s)^2 = \arg\min_{s \in S} d(r,s).$$

Example 6.14. In a binary antipodal signaling scheme, the message S is randomly selected from the alphabet set $\mathcal{S} = \{3, -3\}$ with $p_{-3} = P[S=-3] = 0.3$ and $p_3 = P[S=3] = 0.7$. The message is corrupted by an independent additive noise $N \sim \mathcal{N}(0,2)$. Find the MAP detector $\hat{s}_{\mathrm{MAP}}(r)$ and the corresponding error probability.

(a) (b)

$P(\mathcal{E})$

$\mathcal{N}(m, \sigma^2)$

$p_1 = 0.3 \qquad \sigma = \sqrt{2}$

$p_2 = 0.7 \qquad s_1 = -3$

$s_2 = 3$

(a)

$S = \{s_1, s_2\} = \{-3, 3\}$

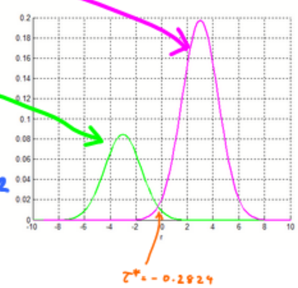
For a given r,

$\hat{s}(r) = s_2$ if

$$\underbrace{p_2}_{P[S=s_2]} f_N(r-s_2) > \underbrace{p_1}_{P[S=s_1]} f_N(r-s_1)$$

$$f_N(n) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{n}{\sigma}\right)^2}$$

$$p_2 \cancel{\frac{1}{\sqrt{2\pi}\,\sigma}} e^{-\frac{1}{2}\left(\frac{r-s_2}{\sigma}\right)^2} > p_1 \cancel{\frac{1}{\sqrt{2\pi}\,\sigma}} e^{-\frac{1}{2}\left(\frac{r-s_1}{\sigma}\right)^2}$$



$$e^{-\frac{1}{2\sigma^2}\left((r-s_2)^2-(r-s_1)^2\right)} > \frac{p_1}{p_2}$$

$$\frac{1}{2\sigma^2}\left((2r-s_2-s_1)(\underbrace{s_2-s_1}_{\text{assume } >0})\right) > \ln\frac{p_1}{p_2}$$

$$r > \underbrace{\frac{\sigma^2 \ln\frac{p_1}{p_2}}{s_2-s_1} + \frac{s_2+s_1}{2}}_{\tau^* \,=\, -0.2824}$$

binary decision threshold

$$\hat{s}_{\mathrm{MAP}}(r) = \begin{cases} s_1, & r \le \tau^*, \\ s_2, & r > \tau^*. \end{cases}$$

(-3) (3) -0.2824

(b) Probability of error

Here, we will try to get $P(\mathcal{E})$ for detector of the form $\hat{s}(r) = \begin{cases} s_1, & r < \tau, \\ s_2, & r \geq \tau. \end{cases}$

Note that our calculation below will work with any τ

(not necessarily the optimal one).

total prob. theorem

$$P(\mathcal{E}) = \underbrace{P(\mathcal{E} \mid S = s_1)}\, p_1 + \underbrace{P(\mathcal{E} \mid S = s_2)}\, p_2$$

$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$

The conditional part of the error event means that this is the error that occur when we assume that $S = s_1$ was transmitted. This error occurs iff the detector detects s_2 (instead of s_1). By the form of the detector above, this occurs iff the received signal $R > \tau$.

$$= P[R \geq \tau \mid S = s_1]$$
$$= P[S + N \geq \tau \mid S = s_1]$$
$$= P[s_1 + N \geq \tau \mid S = s_1]$$
$$= P[N \geq \tau - s_1 \mid S = s_1]$$

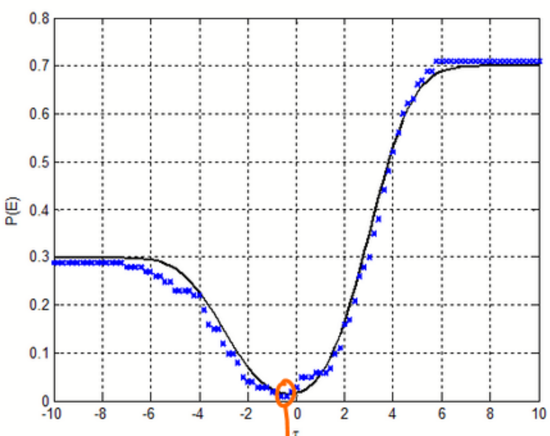
$N \perp\!\!\!\perp S$

$$= P[N \geq \tau - s_1]$$

$$= P[R < \tau \mid S = s_2]$$
$$= P[S + N < \tau \mid S = s_2]$$
$$= P[s_2 + N < \tau \mid S = s_2]$$
$$= P[N < \tau - s_2 \mid S = s_2]$$

$N \perp\!\!\!\perp S$

$$= P[N < \tau - s_2]$$



min $P(\mathcal{E})$

$$P(\mathcal{E}) = p_1 P[N \geq \tau - s_1] + p_2 \underbrace{P[N < \tau - s_2]}_{1 - P[N \geq \tau - s_2]}$$

$N \sim \mathcal{N}(0, \sigma^2)$

$$= p_1 Q\left(\frac{\tau - s_1}{\sigma}\right) + p_2 \Big(\underbrace{1 - Q\left(\frac{\tau - s_2}{\sigma}\right)}_{Q\left(\frac{s_2 - \tau}{\sigma}\right)}\Big)$$

$$= p_1 Q\left(\frac{\tau - s_1}{\sigma}\right) + p_2 Q\left(\frac{s_2 - \tau}{\sigma}\right)$$

For MAP detector, $P(\mathcal{E}) = p_1 Q\left(\frac{\tau^* - s_1}{\sigma}\right) + p_2 Q\left(\frac{s_2 - \tau^*}{\sigma}\right) \approx 0.0153$

For ML detector, $\tau = 0 \Rightarrow P(\mathcal{E}) = 0.0169 \quad (> 0.0153)$

What about exponential noise:

Recall: X is an exponential RV if

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$