Suppose that a Wiener process \((X_t : t \geq 0)\) defined on \((\Omega, \mathcal{F})\) and having drift parameter \(\mu\) and diffusion coefficient \(\sigma^2\) is being continuously observed beginning at \(t=0\). Let the cost of observation per unit time be \(b(>0)\). Assume that the future is discounted continuously at a constant rate \(\delta(>0)\).

The purpose is to test

\[(1) \quad H_0 : \mu = 0 \quad \text{vs} \quad H_1 : \mu = \mu_0 (> 0),\]

given that \(\sigma^2\) is known. If \((W_t : t \geq 0)\) denotes a standard Wiener process, then

\[X_t = \theta \mu_0 t + \sigma W_t \quad (\sigma > 0)\]

and the hypotheses in (1) may be represented as

\[(1)' \quad H_0 : \theta = 0 \quad \text{and} \quad H_1 : \theta = 1.\]

Testing these hypotheses may be carried out by using the continuous time analogue of SPRT given by Dvoretzky, Kiefer and Wolfowitz (1953) for processes with stationary independent increments. The procedure SPRT \((a_0, a_1)\) prescribes stopping according
to the rule
\[(2) \quad \tau^0 = \inf \{ t \geq 0 : X_t \not\in (a_0, a_1) \} \quad (a_0 \leq a_1)\]
and on the set \( \{ \tau^0 = s \} \), it decides in favour of 
\( H_0 \) if \( X_s = a_0 \) and \( H_1 \) if \( X_s = a_1 \). Further, by 
the arguments of the discrete case, it is known that 
the Bayes procedure with constant costs and no 
discounting is the SPRT. Presently, we shall show 
that this result holds in the problem of testing the 
mean of a Wiener process even in a discounted setup. 
It may be noted that in the undiscounted case, 
Dvoretzky et al. have also observed optimality of 
SPRT (see also Krylov and Miroshnichenko, 1980).

Now, consider \( \theta \) as a \((0,1)\)-random variable 
generated by the specified prior probabilities 
\((\pi, 1-\pi)\) on \( (\mathcal{A}, \mathcal{B}) \). This then defines the probability 
measure \( P^\pi \) on the trajectories of the Wiener process 
\( (X_t) \) analogous to the \( P^\pi \) in Chapters 2 and 3. In 
particular,
\[ P^\pi (\theta = 0) = \pi = 1 - P^\pi (\theta = 1). \]

Let \( P_0 \) and \( P_1 \) denote respectively, the probability 
measures governing \( (X_t) \) under \( H_0 \) and \( H_1 \). Thus 
\( P_0 \) and \( P_1 \) are Wiener measures with a common diffusion 
coefficient \( \sigma^2 \) but 0 and \( \mu_0 \) as drift rates respectively,
Define \( \Pi_t = P^\pi (\theta = 0 \mid \mathcal{F}_t) \) where \( \mathcal{F}_t \) is the history of \( (X_t) \) at time \( t \). Then

\[
\pi_t = \pi \frac{dP_0}{dP^\pi}(\mathcal{F}_t)
\]

It is known from Liptser and Shiryaev (1977) that the process \((\bar{W}_t, \mathcal{F}_t)\), where

\[
\bar{W}_t = X_t - \int_0^t \pi_s ds \quad (t \geq 0)
\]

is a square integrable martingale with continuous trajectories and hence a Wiener process, with

\[
\mathbb{E}^\pi \left[ (\bar{W}_t - \bar{W}_s)^2 \mid \mathcal{F}_s \right] = t-s \quad (t \geq s).
\]

Also, the process \((\Pi_t, \mathcal{F}_t)\) with Itô differential

\[
d\Pi_t = \frac{\mu^0}{\sigma} \pi_t (1 - \pi_t) d\bar{W}_t, \quad t \geq 0
\]

and \( \Pi_0 = \pi \) is a strong Markov process under the \( P^\pi \) measure \((0 \leq \pi \leq 1)\). It has drift zero and has diffusion coefficient equal to \( (\mu^0 / \sigma) (1 - \pi) \).

Let \( \sigma = (\tau, \delta) \) denote a sequential decision procedure for the problem in (1). The discounted sampling cost under \( \sigma \) is given by

\[
\int_0^\tau e^{-\delta t} b \, dt = \frac{b}{\delta} (1 - e^{-\delta \tau}).
\]

Assume that the terminal loss function \( l(d, \theta) \) is as
defined in (2) of Chapter 2.

Then, the risk under discounting, say \( R(\sigma ; \beta) \), of \( \sigma \) is

\[
R(\sigma ; \beta) = E^\pi \left[ \frac{d}{d \pi} (1 - e^{-\alpha t}) + e^{-\alpha t} l(a, c) \right].
\]

Let \( \{\sigma\} \) consist of sequential decision procedures \( \sigma \) whose risk under discounting is defined. The Bayes procedure \( \sigma^* = (\tau^*, d^*) \) and its risk \( R^*(\pi; \beta) \) are such that

\[
R^*(\pi; \beta) = R(d^*; \beta) = \inf_{\{\sigma\}} R(\sigma; \beta).
\]

It is the purpose of the present section to show that \( \sigma^* \) is the Wald SPRT.

The continuous time version of Lemma 2.1 may be established on the same lines as the discrete case. Thus the terminal rule under \( \sigma^* \) is

\[
d^* = 1 \quad \text{if} \quad W_0 \tau^* \leq W_1 (1 - \tau^*)
\]
\[
= 0 \quad \text{if} \quad W_0 \tau^* > W_1 (1 - \tau^*)
\]

and the solution of the following stopping problem

\[
\inf_{\tau} E^{\pi} \left[ e^{-\alpha \tau} \left( g(\pi_\tau) - \frac{b}{\beta} \right) + \frac{b}{\beta} \right]
\]
yields \( \tau^* \), where \( g(x) = \min\{w_0 x, w_1 (1-x)\} \) and \( \mathcal{F} \) is the class of Markov times wrt \( (\mathcal{F}_t) \).

Towards solving (10), consider the problem
(11) \[ R^*_1(\pi; \beta) = \inf_{\mathcal{F}} \mathbb{E}^\pi \left[ e^{-\gamma t} g'(\Pi_t) \right] \]
where \( g'(x) = g(x) - \frac{b}{\gamma} \), a bounded continuous function. The Markov process \((\Pi_t)\) and the function \( g' \) satisfy the respective conditions laid down in Grigelionis and Shiryaev (1986) (or Shiryaev, 1978). Solution to (11) may be obtained as the limit of the solution to the version of (11) corresponding to a set of nested discrete sequences of Markov times. Specifically, let \( \mathcal{F}_n \) denote the collection of random variables \( \tau \) which take values in \( \{ k 2^{-n}, k = 0, 1, \ldots \} \) and are such that

\[ \{ \tau \leq k 2^{-n} \} \in \mathcal{F}_n \]
where \( \mathcal{F}_n \) is the \( \sigma \)-field induced by the values of the realizations of \((X_t)\) at time instants \( i 2^{-n}, i = 0, 1, \ldots, 1 \leq k \). Define \( Q_n \) operating on \( g' \) by

\[ Q_n g'(\pi) = \min \left\{ g'(\pi), (e^{-\gamma}) 2^{-n} T_n g'(\pi) \right\} \]

where \( T_n g'(\pi) = \mathbb{E}^\pi \left[ g'(\Pi_t) \right], t \geq 0. \)

**Definition.** A process \((X_t, \mathcal{G}_t)\) is said to be a (continuous time) \((\cdot)\)-submartingale wrt \( \Gamma \) if

1. \( \mathbb{E}[X_t] < \infty \) \( \forall t \geq 0 \), and
2. \( \mathbb{E}(X_t | \mathcal{G}_s) \geq y_{s,t} \beta(t-s) \),
t > s, \beta > 0. It is said to be \textit{maximal} relative
to a measurable function \( h \) if
\( Y_t \leq h(Y_t) \) and
\( Y_t \geq U_t \) for every \((\cdot)\)-submartingale \((U_t, \mathcal{F}_t)\) which
is majorised by \( h \).

Now consider the problem
\[
R_{1,n}^* (\pi; \cdot) = \inf_{\tau} \mathbb{E}^\pi \left[ e^{-\tau t} g' (\Pi \tau) \right].
\]
Then (see (a) - (c) of Section 2.5),
\[
R_{1,n}^* (\pi; \cdot) = \lim_{N \to \infty} Q_n \mathbb{E}^\pi g' (\pi).
\]

It can further be shown on the same lines as Theorems
4 and 5 of Grigelionis and Shiryaev that

(a) \( R^* (\pi; \cdot) = \frac{\beta}{t} + \lim_{n \to \infty} R_{1,n}^* (\pi; \beta) \)

(b) \( (R^* (\Pi_t; \cdot)) \) is the maximal \((\cdot)\)-submartingale
wrt \( \mathbb{F}^\pi \) relative to \( g \) satisfying, for every positive
\( \varepsilon \), sufficiently small,

\[
(12) \quad R^* (\pi; \cdot) = \frac{\beta}{t} + \min \left\{ \begin{array}{l} g' (\pi), \quad e^{-\tau (\pi)} g' (\Pi_t) R^* (\Pi_t; \cdot) \end{array} \right\}
\]

where \( \tau(\varepsilon) \) is the time of first exist from the interval
\((\pi - \varepsilon, \pi + \varepsilon)\), i.e.,
\( \tau(\varepsilon) = \sup \{ t : \pi_s (\pi - \varepsilon, \pi + \varepsilon) \},
\]
s \( \leq t \) . It is clear from (12) that the \( \mathbb{F}^\pi \)-time
\[
\tau_0 = \inf \{ t \geq 0 : R^* (\pi_t; \cdot) = \varepsilon (\pi_t) \}.
\]
is optimal for (10) if \( P^\tau ( \tau_0 < \omega ) = 1 \). The latter, and the fact that \( \tau_0 \) is actually the stopping time of an SPRT will be shown in the following.

Concavity of \( g(\cdot) \) and \( R^*(\cdot; \beta) \) continue to hold for the same reasons as in the discrete case (see Lemma 2.2 and Theorem 2.3). It follows from those facts that the region \( C(\tau_0) \) of continuation at any stage under \( \tau_0 \) is a compact set. In terms of \( \pi \), it is an interval, i.e., there exist constants \( A_0 \) and \( A_1 \) independently of the prior probability \( \pi \), such that at any stage

\[
C(\tau_0) = \{ \omega : A_0 < \Pi_t(\omega) < A_1 \}
\]

In view of (4) and (5), we may rewrite \( C(\tau_0) \) as

\[
C(\tau_0) = \{ \omega : a_0 < X_t(\omega) < a_1, \ t \geq 0 \}
\]

Since the probability is one that the Wiener process \( X_t \) will eventually move out of the interval \( (a_0, a_1) \), \(-\infty < a_0 < a_1 < \infty\), the Markov time \( \tau_0 \) is finite (a.s.). Thus, the optimal stopping time for (10) is \( \tau_0 \), i.e., \( \tau^* = \tau_0 \) and the Bayes procedure for the problem in (8) is the SPRT.
4.2 Minimax Procedure Under Discounting.

Consider the problem of testing for the sign of \( \mu \), the drift rate of a Wiener process \( (X_t) \) having a known diffusion coefficient, i.e., the problem specified by

\[
H_0 : \mu \leq 0 \quad \text{vs} \quad H_1 : \mu > 0.
\]

Let the loss function be

\[
l(d, \mu) = \begin{cases} 
  c \mu^r & \text{if } d=1 \text{ and } \mu \leq 0 \\
  c\mu^r & \text{if } d=0 \text{ and } \mu > 0 \\
  0 & \text{otherwise.}
\end{cases}
\]

DeGroot (1960) has established that under this loss function with \( 0 < r < 2 \) the minimax sequential procedure for (13) is a symmetric SHRT. We shall prove that this property holds for \( 0 < r < 1 \) when the future is continuously discounted.

Employing the notation of Section 4.1, and the loss function (14) with \( 0 < r < 1 \), we see that as in (2) the risk of a sequential decision procedure \( \sigma \) will be

\[
E^\mu \left[ \frac{1}{\nu} (1 - e^{-\tau \nu}) + e^{-\tau \nu} l(d, \mu) \right]
\]

where \( E^\mu \) denotes expectation under drift rate \( \mu \).
The risk, after ignoring a constant term, wrt which we seek the minimax procedure, say $\delta^*$, is

\begin{equation}
R(\delta^*, \mu) = \mathbb{E}^\mu \left[ e^{-q \tau} \ell(d, \mu) - \frac{h}{\varepsilon} e^{-q \tau} \right].
\end{equation}

Procedure $\delta^*$ is by definition such that

\begin{equation}
\sup \frac{R(\delta^*, \mu)}{\delta} = \inf \sup \frac{R(\delta, \mu)}{\mu}
\end{equation}

Consider first the class $\delta_h$ of symmetric SHRT's, SHRT $(-\frac{h}{2}, \frac{h}{2})$ for various $h > 0$. That is, under $\delta_h$, $(X_t)$ is observed only as long as $|X_t| < \frac{h}{2}$ and a decision favouring $H_0(H_1)$ is taken when $X_t \leq -\frac{h}{2}$ ($X_t \geq \frac{h}{2}$). If $\delta_h^*$ denotes the minimax rule in the class $\delta_h$, then $\delta_h^*$ would be determined if one finds a pair $(\mu^*, h^*)$ such that

\begin{equation}
R(\delta_h^*, \mu^*) = \inf \frac{R(\delta_h, \mu^*)}{\mu} = \sup \frac{R(\delta_h^*, \mu)}{\mu}
\end{equation}

It will next be shown, as in DeGroot (1960), that $\delta_h^*$ is in fact $\delta^*$, the procedure minimax in the class $\{\delta\}$. Existence of $\delta_h^*$ and $\delta^*$ is guaranteed by the fact that in each case the action space $\{0, 1\}$ is finite and the risk set is closed and bounded.

Let $\tau$ denote the time of the first passage of $(X_t)$ outside of $(-\frac{h}{2}, \frac{h}{2})$. Then, by assuming $\sigma = 1$ and knowing the distribution of $\tau$
(see e.g., Darling and Sigort, 1953) we get

\[ R''(e^{-\beta t}) = e^{\beta h/2} \left( \frac{2}{2 \cosh (\frac{h}{2} \sqrt{\mu^2 + 2\beta})} \right) \]

and

\[ R''(e^{-\beta t}) = \frac{\cosh (\mu h/2)}{\cosh (\frac{h}{2} \sqrt{\mu^2 + 2\beta})} \]

Using these expressions in (15),

\[ R(\sigma_h, \mu) = \frac{e^{-\mu h/2} - 2(h/\mu) \cosh(\mu h/2) }{2 \cosh (\frac{h}{2} \sqrt{\mu^2 + 2\beta})} \]

\[ = \frac{c(\mu)^{1/r} \cosh(\mu h/2) - 2(h/\mu) \cosh(\mu h/2)}{2 \cosh (\frac{h}{2} \sqrt{\mu^2 + 2\beta})} \]

Since \( R(\sigma_h, \mu) = R(\sigma_h, -\mu) \), it is sufficient to consider only positive \( \mu \). Our concern will thus be only with the first expression in (18). We make the following transformations.

\[ a = \left( \frac{2h}{1c} \right)^{1/r} \]

\[ m = \left( \frac{2a}{2e} \right)^{1/r} \]

\[ \rho = c \left( \frac{2a}{2e} \right)^{2/r} \]

Then,

\[ R^1(a, m) = \frac{2m^r - \frac{m^a}{2} - 1}{2 \frac{m^a}{2} \cosh (a\theta/2)} \]

where \( \theta = \sqrt{m^2 + 2\rho} \), is the function we will be
concerned with instead of \( R(\delta_1, \mu) \).

We fix \( a(>0) \) and obtain \( m_a^* \), the value of \( m \) which corresponds to the maximum of \( R^1(a, m) \). As \( a \) is varied, \( R^1(a, m_a^*) \) varies. Let \( a_m^* \) correspond to the maximum of \( R^1(a, m_a) \). Then, \((a_m^*, m^*)\) is a saddle point and \( a_b^* \) would be the minimax sequential procedure. In what follows, we show that (I) a unique \( m \) maximises \( R^1 \) for every fixed \( a \) and (II) a unique \( a \) minimises \( R^1 \) for every fixed \( m \).

For a fixed \( a \), \( \frac{\partial R^1}{\partial m} = 0 \) gives,

\[
0 = -\frac{2m^r - e^m}{2e^m\cosh^2(a\theta/2)} \cdot \frac{e^m}{2} \cdot (\frac{m}{\theta} \tanh(\frac{\theta}{2}) + \cosh(\frac{\theta}{2}) ) + \frac{2r m^{r-1} - e^m}{2e^m\cosh(a\theta/2)} \cdot \frac{e^m}{2}.
\]

That is,

\[(20) \quad 0 = \frac{\theta}{2} (2m^r - e^m - 1)(\frac{m}{\theta} \tanh(\frac{\theta}{2}) + 1 ) + 2r m^{r-1} - e^m \cdot \frac{e^m}{2}.
\]

We rearrange terms in (20) to get

\[(21) \quad 2rm^{r-1} = e^m (1 - \frac{m}{\theta} \tanh(\frac{\theta}{2}) )
\]

\[ + \frac{\theta}{2} (2m^r - 1)(1 + \frac{m}{\theta} \tanh(\frac{\theta}{2}) )\.]
Denote the LHS and RHS of (21) by \( \phi_a(m) \) and \( \psi_a(m) \) respectively. Now note that for every fixed \( a > 0 \),

(i) \( \psi_a \) is continuous and increasing;

(ii) \( \psi_a \to \frac{a}{2} \) as \( m \to 0 \) and \( \to \infty \) as \( m \to \infty \);

(iii) \( \phi_a \) is continuous and, for \( r < 1 \), decreasing;

(iv) \( \phi_a \to +\infty \) as \( m \to 0 \), and \( \to 0 \) as \( m \to \infty \).

These observations establish that there is a unique \( m \), say \( m_a \), for which (21) will be satisfied under the condition that \( r < 1 \). We may also note that

\[
\frac{R^1}{m_a} = \left( 2 \alpha \frac{m_a}{2} \cosh \frac{ma}{2} \right) \left( \phi_a(m) - \psi_a(m) \right)
\]

is negative at \( m_a \), so that \( m_a \) corresponds to the maximum of \( R^1 \).

\[
\text{II. } \frac{dR^1}{da} = 0 \text{ gives, for fixed } m,
\]

\[
0 = - \frac{2m^r - \frac{ma}{2} - 1}{2 \alpha m_a - \cosh (a0/2)} \left( \frac{2}{2} \sinh \frac{ma}{2} + \frac{ma}{2} \cosh \frac{ma}{2} \right)
\]

\[
- \frac{ma}{2 \alpha m_a/2 \cosh (a0/2)}
\]
Let us restate this equation as

\[(23) \quad e^{mb} \left( \frac{m}{2} - \frac{g}{2} \tanh \frac{\theta}{2} \right) = (1-2m^r) \left( \frac{m}{2} + \frac{g}{2} \tanh \frac{\theta}{2} \right)\]

and denote by \( \phi_m(a) \) and \( \psi_m(a) \) the LHS and the RHS respectively.

Again \( \phi_m(.) \) and \( \psi_m(.) \) are continuous functions with

\[\phi_m(a) \to \frac{m}{2} \quad \text{as} \quad a \to 0, \quad \text{and} \quad \to -\infty \quad \text{as} \quad a \to \infty,\]

\[\psi_m(a) \to (1-2m^r) \frac{m}{2} \quad \text{as} \quad a \to 0, \quad \text{and} \quad \to (1-2m^r) \left( \frac{m}{2} + \frac{g}{2} \right) \quad \text{as} \quad a \to \infty.\]

Since \( \lim_{a \to 0} \phi_m(a) > \lim_{a \to 0} \psi_m(a) \) for every \( m \), in each of the three cases, namely, \((1-2m^r) > 0, = 0, < 0\) \( \phi_m \) and \( \psi_m \) intersect at only one positive \( a \). Thus (23) yields a unique \( a \), say \( a_m \), for every fixed \( m \) and it is such that

\[(24) \quad \frac{1}{\delta_a} = (2e^{\frac{mb}{2}} \cosh \frac{\theta}{2}) (\psi_m(a) - \Phi_m(a))\]

\[< 0 \quad \text{on} \quad (0, a_m)\]

\[= 0 \quad \text{on} \quad \{a_m\}\]

\[> 0 \quad \text{on} \quad (a_m, \infty),\]

implying that \( a_m \) corresponds to the minimum.
Simultaneous solution of equations (21) and (23) yields the required point in the \((a, m)\)-space, namely, \((a^*, m^*)\) from which one gets, using (19), the minimax SRT \(a^*_h\). Explicit solutions for \(a\) and \(m\) from (21) and (23) are not obtained. However the solution does correspond to the minimax procedure since it can be shown to be the saddle point of \(a^*_h\).

\[(a^*, m^*)\] is a saddle point if

\[\alpha \beta - \gamma^2 < 0 \quad \text{at} \quad (a^*, m^*)\]

where \[\alpha = \frac{2R_1}{\varphi^2}, \quad \beta = \frac{2R_1}{\varphi m^2}, \quad \gamma = \frac{2R_1}{\varphi m}\]; in the light of (22) and (24), it is clear that the solution of (21) and (23) yields the saddle point.

Arguments identical to DoGroot's may now be advanced to show that \(a^*_h\) is in fact \(a^*\).

By the result of Section 4.1, the Bayes procedure for

\[H_0^I: \mu = \mu_0 \quad \text{vs} \quad H_1^I: \mu = -\mu_0\]

is a symmetric SRT and this has to be \(a^*_h\).

Application of Theorem 5.3 of Berger (1980) completes the proof.
4.3 Conclusion.

The problem of testing for the sign of a Wiener process mean $\mu$ is also discussed in Chernoff (1968) wherein the Bayes procedure with respect to a normal prior for $\mu$ under the loss function in (14) with $r=1$ is characterised. The loss may in fact be taken to be any integrable function of $(s, Y(s))$ where $Y(s)$ is the state of the posterior process (itself a Wiener process) at 'time' $s$ (a function of $t$ and $\sigma$). Similar approach is clearly possible in our problem in Section 4.1: $(t, \Pi_t)$ is a Markov process wrt $(\mathcal{F}_t)$ and the loss at time $t$ is a fixed function of $(t, \Pi_t)$ and the problem is one of optimally stopping the Markov process $(t, \Pi_t)$, so that the theory in Grigolionis and Shiryaov (1966) may be applied directly. But it would not be apparent by this approach that the continuation region is the same in successive stages, due to nonhomogeneity of the process $(t, \Pi_t)$. On the other hand, explicit use of the manner of dependence of the loss on $t$, as in Section 4.1, leads to the conclusion that the discount optimal procedure for testing the Wiener mean is the SRT. It is also shown in the foregoing that the
discount-optimal procedure is also minimax wrt the loss in (14) for $0 < r < 1$. This establishes as in the undiscounted case, a degree of robustness of the Bayes solution of the problem even under discounting.