SOME EXPONENTIAL DISTRIBUTIONS
ASSOCIATED WITH BROWNIAN MOTION

by

Nils Lid Hjort
Some exponential distributions associated with Brownian motion

Nils Lid Hjort

University of Oslo and Norwegian Computing Centre

-- March 1991 --

ABSTRACT. Let \( V_\varepsilon(b) \) be \( 1/\varepsilon \) times the total time a Brownian motion \( W(s) \) spends in the strip \([bs - \frac{1}{2}\varepsilon, bs + \frac{1}{2}\varepsilon]\). It is shown to converge to a process \( V(b) \) with exponentially distributed marginals, and \( V(b) \) is simply exponential with parameter \(|b|\). This construction leads to families of bivariate and multivariate exponential distributions. Truncated versions of such 'total relative time' variables are also studied. A relation is pointed out to a second order asymptotics problem in statistical estimation theory, recently investigated in Hjort and Fenstad (1991a, 1991b).

KEY WORDS: Brownian motion, exponential process, multivariate exponential distribution, second order asymptotics for estimators, total relative time

1. Introduction and summary. For a standard Brownian motion process \( \{W(s): s \geq 0\} \), consider the total amount of time where \( W(s) \) visits the strip \([bs - \frac{1}{2}\varepsilon, bs + \frac{1}{2}\varepsilon]\), divided by \( \varepsilon \), i.e.

\[
V_\varepsilon(b) = \frac{1}{\varepsilon} \mu\{s \geq 0: bs - \frac{1}{2}\varepsilon \leq W(s) \leq bs + \frac{1}{2}\varepsilon\} = \frac{1}{\varepsilon} \int_0^\infty I\{bs - \frac{1}{2}\varepsilon \leq W(s) \leq bs + \frac{1}{2}\varepsilon\} \, ds,
\]

writing \( \mu \) for Lebesgue measure on the halflife and \( I\{\ldots\} \) for indicator functions. It is shown in Section 2 that \( V_\varepsilon(b) \) converges in distribution to a certain \( V(b) \), which we may regard as the total relative time spent by \( W(s) \) along the ray \( w = bs \). It turns out that \( V(b) \) is simply exponentially distributed with parameter \(|b|\). There is also process convergence of \( V_\varepsilon(\cdot) \) to \( V(\cdot) \), where \( \{V(b): b \neq 0\} \) is a process with continuous sample paths and with exponentially distributed marginals. Its covariance and correlation structure is found in Section 3. In particular this construction leads to families of bivariate and multivariate exponential distributions.

In Section 4 a more general variable \( V(a,b) \) is studied, defined as the total relative time during \([a, \infty)\) that \( W(s) \) spends along \( bs \). The distribution of \( V(a,b) \) is a mixture of an exponential and a unit point mass at zero. A simple consequence of this result is a rederivation of a well known formula for the distribution of the maximum of Brownian motion over an interval. Finally some supplementing results and remarks are given in Section 5. In particular some consequences for empirical partial sum processes are briefly discussed.

A certain second order asymptotics problem in statistical estimation theory led by serendipity to the present study on total relative time variables for Brownian motion. Suppose \( \{\theta_n: n \geq 1\} \) is an estimator sequence for a parameter \( \theta \), where \( \theta_n \) is based on
the first \( n \) data points in an i.i.d. sequence, and consider \( Q_\delta \), the number of times, among \( n \geq a/\delta^2 \), where \( |\theta_n - \theta| \geq \delta \). Almost sure convergence (or strong consistency) of \( \theta_n \) is equivalent to saying that \( Q_\delta \) is almost surely finite for every \( \delta \), and it is natural to inquire about its size. A particular result of Hjort and Fenstad (1991a, Section 7) is that under natural conditions, which include the existence of a normal \((0, \sigma^2)\) limit for \( \sqrt{n}(\theta_n - \theta) \),

\[
\delta^2 Q_\delta \to_d Q = Q(a,1/\sigma) = \int_a^\infty I\{|W(s)| \geq s/\sigma\} ds
\]

as \( \delta \to 0 \). If \( \{\theta_{n,1}\} \) and \( \{\theta_{n,2}\} \) are first order equivalent estimator sequences, with the same \( N(0, \sigma^2) \) limit for \( \sqrt{n}(\theta_{n,j} - \theta) \), and \( Q_{\delta,j} \) is the number of \( \delta \)-misses for sequence \( j \), then \( Q_{\delta,1}/Q_{\delta,2} \to 1 \) and \( \delta^2(Q_{\delta,1} - Q_{\delta,2}) \to 0 \) in probability. One way of distinguishing between the two estimation methods is by studying second order aspects of \( Q_{\delta,1} - Q_{\delta,2} \). It turns out that \( \delta \) times the difference in typical cases has a limit distribution which is a constant times \( V(a,1/\sqrt{\sigma}) - V(-1/\sqrt{\sigma}) \) or times the simpler \( V(1/\sqrt{\sigma}) - V(-1/\sqrt{\sigma}) \) if \( a = a(\delta) \) is allowed to decrease to zero in the definition of \( Q_{\delta,j} \). Note the connection from \( Q(a,1/\sigma) \) of (1.2) to \( V(a,\pm 1/\sigma) \). Some further details are in 5C in the present paper, while further background and discussion can be found in Hjort and Fenstad (1991b).

2. The total relative time along a ray. Let \( b \neq 0 \), and consider \( V_\varepsilon(b) \) as defined in (1.1). Note that the integral is a.s. finite since \( W(s)/s \) goes a.s. to zero when \( s \) grows. If \( c \) is positive, write \( V \sim \text{Exp}(c) \) for the exponential distribution with density \( g(v) = ce^{-cv} \) for \( v \geq 0 \). We will prove convergence in distribution

\[
V_\varepsilon(b) \to_d V(b) \sim \text{Exp(|b|)}
\]

by demonstrating appropriate convergence of all moments. This is sufficient since the exponential distribution is determined by its moment sequence.

For the first moment, observe that

\[
EV_\varepsilon(b) = \frac{1}{\varepsilon} \int_0^\infty \Pr\{bs - \frac{1}{2}\varepsilon \leq W(s) \leq bs + \frac{1}{2}\varepsilon\} ds \approx \int_0^\infty f_s(bs) ds + O(\varepsilon),
\]

where \( f_s(x) = \phi(x/\sqrt{s})/\sqrt{s} \) is the density function for \( W(s) \). Accordingly \( EV_\varepsilon(b) \) goes to \( \int_0^\infty \phi(b\sqrt{s})/\sqrt{s} ds = 1/|b| \). Next consider the \( p \)-th moment. One finds

\[
EV_\varepsilon(b)^p = \frac{1}{\varepsilon^p} \int_0^\infty \cdots \int_0^\infty \Pr\{W(s_1) \in bs_1 \pm \frac{1}{2}\varepsilon, \ldots, W(s_p) \in bs_p \pm \frac{1}{2}\varepsilon\} ds_1 \cdots ds_p
\]

\[
= \frac{1}{p!} \int_0^\infty \cdots \int_{s_1 < \cdots < s_p} f_{s_1,\ldots,s_p}(bs_1, \ldots, bs_p) ds_1 \cdots ds_p + O(\varepsilon),
\]

where \( f_{s_1,\ldots,s_p}(x_1, \ldots, x_p) \) is the density function of \( (W(s_1), \ldots, W(s_p)) \). By the Gaussian and Markovian properties of \( W(.) \) this density can in fact be written

\[
\phi\left(\frac{x_1}{\sqrt{s_1}}\right) \frac{1}{\sqrt{s_1}} \phi\left(\frac{x_2 - x_1}{\sqrt{s_2 - s_1}}\right) \frac{1}{\sqrt{s_2 - s_1}} \cdots \phi\left(\frac{x_p - x_{p-1}}{\sqrt{s_p - s_{p-1}}}\right) \frac{1}{\sqrt{s_p - s_{p-1}}}
\]

(2.3)
when $s_1 < \cdots < s_p$. To carry out the $p$-dimensional integration, insert $(b s_1, \ldots, b s_p)$ for $(x_1, \ldots, x_p)$, and transform to new variables $u_1 = s_1, u_i = s_i - s_{i-1}$ for $i = 2, \ldots, p$. The result is then that

$$EV_{\epsilon}(b)^p \rightarrow p! \int_0^\infty \cdots \int_0^\infty \frac{\phi(b \sqrt{u_1})}{\sqrt{u_1}} \cdots \frac{\phi(b \sqrt{u_p})}{\sqrt{u_p}} du_1 \cdots du_p = p!(1/|b|)^p$$

for each $p$. But this is manifestly the moment sequence of $\text{Exp}(|b|)$, proving (2.1).

The case of $b = 0$ is different, since $W$ spends an infinite amount of time along the time axis. An interesting $|N(0,1)|$ limit result for the relative time in $\pm \epsilon$ during $[0,T]$ is in 5A below.

3. The exponential process. We have seen that $V_{\epsilon}(b)$ goes to an exponentially distributed $V(b)$, and in the same manner we should find bivariate and multivariate exponential distributions by considering two or more $b$'s at the same time. This requires verification of simultaneous convergence in distribution of $(V_{\epsilon}(b_1), \ldots, V_{\epsilon}(b_n))$ and similar quantities. This section indeed demonstrates process convergence of $V_{\epsilon}(\cdot)$ to $V(\cdot)$, and studies some of the properties of the limiting process.

3A. Process convergence. The first main result is as follows.

THEOREM. There is a well-defined stochastic process $V = \{V(b) : b \neq 0\}$ with exponentially distributed marginals and with the property that $(V(b_1), \ldots, V(b_n))$ is the limit in distribution of $(V_{\epsilon}(b_1), \ldots, V_{\epsilon}(b_n))$ for each finite set of non-null indexes $b_i$. There exists a version of $V$ with continuous paths, and $V_{\epsilon}(\cdot)$ converges to $V(\cdot)$ in the uniform topology on the $C$-space $C[b_0, b_1]$ of continuous functions on $[b_0, b_1]$, for each interval not containing zero.

PROOF: Consider two rays $b_1$ and $c_1$ and their associated total relative time variables $(V_{\epsilon}(b_1), V_{\epsilon}(c_1))$. Using the Cramér–Wold theorem in conjunction with the moment convergence method we see that convergence of $EV_{\epsilon}(b)^p V_{\epsilon}(c)^q$ to the appropriate limit, for each $p$ and $q$, is sufficient. But this can be proved by slight elaborations on the techniques of Section 2. By Fubini's theorem

$$V_{\epsilon}(b)^p V_{\epsilon}(c)^q = \frac{1}{\epsilon^{p+q}} \int_0^\infty \cdots \int_0^\infty I\{W(s_1) \in bs_1 \pm \frac{1}{2} \epsilon, \ldots, W(s_p) \in bs_p \pm \frac{1}{2} \epsilon, \ldots, W(t_1) \in ct_1 \pm \frac{1}{2} \epsilon, \ldots, W(t_q) \in ct_q \pm \frac{1}{2} \epsilon\} ds_1 \cdots ds_q,$$

and its expected value is seen to converge to

$$EV(b)^p V(c)^q = \int_0^\infty \cdots \int_0^\infty f(s_1, \ldots, s_p, t_1, \ldots, t_q)(bs_1, \ldots, bs_p, ct_1, \ldots, ct_q) ds_1 \cdots dt_q \quad (3.1)$$

by Lebesgue's theorem on dominated convergence. Note that the integral is over all of $[0, \infty)^{p+q}$ and that a simple expression like (2.3) for the density of $(W(s_1), \ldots, W(t_q))$ is only valid when the time-points are ordered, so the factual integration in (3.1) is difficult to carry through (but possible, see 3C below). What is important at the moment is however the mere existence of this and all other similar limits of product moments for the $V_{\epsilon}(\cdot)$-process. We may conclude that all finite-dimensional distributions converge to well-defined
limits. That these finite-dimensional distributions also constitute a Kolmogorov-consistent system is a by-product of the tightness condition verified below.

The $V_\varepsilon(.)$-process has continuous paths in $b \neq 0$ for each $\varepsilon$, since $W(.)$ is continuous. In order to prove process convergence on $C[b_0,b_1]$ for a given interval we need to demonstrate tightness of the $\{V_\varepsilon(.)\}$ family as $\varepsilon$ goes to zero. Note that if $V_\varepsilon^*(b) = V_\varepsilon(-b)$, then the processes $V_\varepsilon^*(.)$ and $V_\varepsilon(.)$ have identical distributional characteristics, so it suffices to consider the positive part of the process. Following results in Shorack and Wellner (1986, page 52) it is enough to verify that

$$\limsup_{\varepsilon \to 0} E\{V_\varepsilon(b+h) - V_\varepsilon(b)\}^4 \leq K h^2 \quad \text{for some } K, \quad (3.2)$$

for all $h \geq 0$ and for all $b$ with $b$ and $b+h$ in $[b_0,b_1]$, where $0 < b_0 < b_1$. By the arguments for finite-dimensional convergence used above the left hand side of (3.2) is equal to $m_b(h) = E\{V(b+h) - V(b)\}^4$. This is seen to be a smooth function of $h$ with finite derivatives at zero. Ingenious and rather elaborate Taylor expansion arguments can in fact be furnished to prove that

$$EV(b+h)^4 = (24/b^4)\{1 - 4\delta + O(\delta^2)\},$$

$$EV(b)V(b+h)^3 = (24/b^4)\{1 - 6\delta + O(\delta^2)\},$$

$$EV(b)^2V(b+h)^2 = (24/b^4)\{1 - 6\delta + O(\delta^2)\},$$

$$EV(b)^3V(b+h) = (24/b^4)\{1 - 6\delta + O(\delta^2)\},$$

$$EV(b)^4 = 24/b^4,$$

where $\delta = h/b$, so that $m_b(h) = K_2(b)h^2 + K_3(b)h^3 + \ldots$, for local constants $K_j(b)$ that are continuous functions of $b$ (as long as $b \neq 0$). This is dominated by a common $Kh^2$ for all $b$ and $b+h$ in the interval under consideration. This verifies (3.2), and incidentally at the same time verifies the so-called Kolmogorov condition for almost sure continuity of the sample paths, see Shorack and Wellner (1986, Chapter 2, Section 3).

Using the moment formula in 3C below one may in fact calculate the left hand side of (3.2) explicitly, and a fair amount of analysis leads to $m_b(h) = 24 \cdot 352 h^2/b^6 + O(h^3)$. The proof above circumvented the need for information on this level of detail, however.

3B. Dependence structure. In order to investigate this to some extent we calculate covariances and correlations. Let $0 < b < c$ and $-c < 0 < d$. Then

$$\text{cov}\{V(b), V(c)\} = \frac{1}{c} \cdot \frac{1}{2c - b} \quad \text{and} \quad \text{cov}\{V(-c), V(d)\} = \frac{1}{d} \cdot \frac{1}{2c + d} + \frac{1}{c} \cdot \frac{1}{2c + d} - \frac{1}{cd}. \quad (3.3)$$

To prove this, consider the case of two positive parameters. Then by previous arguments

$$EV_\varepsilon(b)V_\varepsilon(c) = \int_0^\infty \int_0^\infty \Pr\{W(s) \in bs \pm \frac{1}{2}\varepsilon, W(t) \in ct \pm \frac{1}{2}\varepsilon\} \, dsdt/\varepsilon^2$$

$$\to \int \int_{s < t} \left[f_{s,t}(bs,ct) + f_{s,t}(cs, bt)\right] \, dsdt$$

$$= \int \int_{s < t} \left[\phi(b\sqrt{s})\phi(\frac{ct - bs}{\sqrt{t - s}}) + \phi(c\sqrt{s})\phi(\frac{bt - cs}{\sqrt{t - s}})\right] \frac{1}{\sqrt{s}} \frac{1}{\sqrt{t - s}} \, dsdt,$$
where (2.3) is used again. Now transform first to \((s,u) = (s,t-s)\) and then to \((x,y) = (\sqrt{s}, \sqrt{u})\), to get

\[
4 \int_0^\infty \int_0^\infty \left[ \phi(bx) \phi\left( \frac{cy^2 + (c-b)x^2}{y} \right) + \phi(cx) \phi\left( \frac{by^2 - (c-b)x^2}{y} \right) \right] \, dx \, dy.
\]

The rest of the calculation is carried out using the formula \(\int_0^\infty \exp\left\{-\frac{1}{2} (k^2 y^2 + l^2 y^2)^2 \right\} dy = \frac{1}{2} \sqrt{2\pi} (1/k) \exp(-kl)\). This formula can be proved by clever but elementary integrations, and is valid for positive \(k\) and \(l\). One finds

\[
\frac{4 \sqrt{2\pi}}{2} \int_0^\infty \left[ \frac{1}{c} \exp\left\{-\frac{1}{2} (b^2 + 4c(c-b)x^2) \right\} + \frac{1}{b} \exp\left\{-\frac{1}{2} c^2 x^2 \right\} \right] dx = \frac{1}{c} \frac{1}{2c - b} + \frac{1}{b} c.
\]

The first formula in (3.3) follows from this, and the other case is handled similarly. □

It is convenient to give formulae (3.3) in another form, using \((b,c) = (b,b+h)\) in the first case and \((-c,d) = (-c,kc)\) in the second. Then

\[
\text{cov}\{V(b), V(b+h)\} = \frac{1}{b+h} \frac{1}{b+2h} \quad \text{and} \quad \text{cov}\{V(-c), V(kc)\} = \frac{1}{c^2 (k+2)(2k+1)},
\]

and the correlation coefficients become

\[
\text{corr}\{V(b), V(b+h)\} = \frac{b}{b+2h} \quad \text{and} \quad \text{corr}\{V(-c), V(kc)\} = -\frac{3k}{(k+2)(2k+1)}. \tag{3.4}
\]

For small \(h\) it is worth noting that

\[
E\{V(b+h) - V(b)\} = \frac{1}{b+h} - \frac{1}{b} = -\frac{1}{b^2} h,
\]

\[
E\{V(b+h) - V(b)\}^2 = \frac{2}{b^2} + \frac{2}{(b+h)^2} - \frac{4}{b(b+2h)} = \frac{4}{b^3} h.
\]

3C. Bivariate and multivariate exponential distributions. We have constructed a full exponential process, and in particular \((V(b_1), \ldots, V(b_n))\) is a random vector with dependent and exponential marginals. These bivariate and multivariate exponential classes of distributions appear to be new.

Formula (3.4) shows that if values \(\mu_1 > 0, \mu_2 > 0, \rho \in (0,1)\) are given, then a pair of dependent exponentials \((V(b_1), V(b_2))\) can be found with \(EV(b_1) = \mu_1, EV(b_2) = \mu_2,\) and correlation \(\rho\). The class of bivariate exponential distributions is accordingly rich in the sense of achieving all positive correlations. The negative correlation in (3.4) starts out at zero for \(k\) small, decreases to \(-\frac{1}{3}\) for \(k = 1\), and then climbs up towards zero again when \(k\) grows, so negative correlations between \(-\frac{1}{3}\) and \(-1\) cannot be attained. Note that the maximal negative correlation occurs between \(V(b)\) and \(V(-b)\).

In order to study the bivariate distribution for \((V(b), V(c))\) we calculate its double moment sequence (3.1) explicitly, for the case of \(0 < b < c\). The technique is to split the integral into \(n! = (p+q)!\) parts, corresponding to all different orderings of the \(n = p+q\) time indexes, the point being that a formula like (2.3) for the density of \((W(s_1), \ldots, W(s_n))\)
can be exploited for each given ordering. These orderings can be grouped into \( (n)_p \) types of paths, say \( (e_1s_1, \ldots, e_ns_n) \) where \( s_1 < \cdots < s_n \) and \( e_j \) is equal to \( b \) in exactly \( p \) cases and equal to \( c \) in exactly \( q \) cases. There are \( p!q! \) different paths for given locations for the \( p \) \( b \)'s and \( q \) \( c \)'s, so the full integral can be written \( \sum p!q! g(\text{path}) \), where the sum is over all \( (n)_p \) classes of paths and \( g(\text{path}) \) is the contribution for a specific path of the appropriate type. It remains to calculate the \( g \)-terms of various types, i.e. to evaluate

\[
\int \cdots \int_{0 < s_1 < \cdots < s_n} f_{s_1, \ldots, s_n}(e_1s_1, \ldots, e_ns_n) \, ds_1 \cdots ds_n
\]

for a path with \( e_j \)'s equal to \( b \) or \( c \). Stameniforous integrations, similar to but more strenuous than those used to prove (3.3), show in the end that

\[
g(\text{path}) = \left( \frac{1}{b_0} \right)^{i(0)} \left( \frac{1}{b_1} \right)^{i(1)} \cdots \left( \frac{1}{b_n} \right)^{i(n)},
\]

where the path when read backwards, i.e. looking through \( (e_n, \ldots, e_1) \) in the notation above, has \( i(0) \) \( b \)'s first, then \( i(1) \) \( c \)'s, then \( i(2) \) \( b \)'s, \&cetera. Furthermore \( b_0 = b, b_1 = c \), and \( b_j = b + j(c - b) \). Note that \( i(0) + i(2) + \cdots = p \), \( i(1) + i(3) + \cdots = q \), and \( i(0) + i(1) + \cdots + i(n) = n \). And \( EV(b)PV(c)^q \) is equal to \( p!q! \) times the sum of all such \( g(\text{path}) \) terms.

To illustrate this somewhat cryptic formula, try \( EV(b)^2V(c)^2 \). There are \( \binom{4}{2} = 6 \) types of paths, corresponding to \( (b, b, c, c), (b, c, b, c), (b, c, c, b), (c, b, b, c), (c, b, c, b), (c, c, b, b) \), and each of these has weight \( 2!2! = 4 \). Their \( (i(0), i(1), \ldots, i(4)) \) representations are respectively \( (0, 2, 2, 0, 0), (0, 1, 1, 1, 1), (1, 2, 1, 0, 0), (0, 1, 2, 1, 0), (1, 1, 1, 1, 0), (2, 2, 0, 0, 0) \). Accordingly

\[
EV(b)^2V(c)^2 = 4 \left\{ \frac{1}{b_1 b_2^2} + \frac{1}{b_1 b_2 b_3} + \frac{1}{b_0 b_1 b_2} + \frac{1}{b_0 b_1 b_2} + \frac{1}{b_0 b_1 b_2} \right\},
\]

where \( b_0 = b, b_1 = c, \ldots, b_4 = b + 4(c - b) \).

I have not been able to produce an explicit formula for the joint probability density of \( (V(b), V(c)) \), but at least an expression can be found for its joint moment generating function. It becomes

\[
E \exp\{sV(b) + tV(c)\} = \sum_{n \geq 0} \frac{1}{n!} \sum_{p+q=n} \binom{n}{p} s^p t^q E\{V(b)^p V(c)^q\}
= \sum_{n \geq 0} \frac{1}{n!} \sum_{p+q=n} \frac{n!}{p!q!} s^p t^q \sum_{\text{paths}} p!q! g(\text{path})
= \sum_{p \geq 0, q \geq 0} s^p t^q \sum_{i(0), i(1), \ldots, i(p+q)} \left( \frac{1}{b_0} \right)^{i(0)} \left( \frac{1}{b_1} \right)^{i(1)} \cdots \left( \frac{1}{b_{p+q}} \right)^{i(p+q)},
\]

where again the inner sum is over all \( (p + q)!/p!q! \) types of paths with \( p \) \( b \)'s and \( q \) \( c \)'s, and the multiplicities \( i(0), i(1), \ldots, i(p+q) \) have even-sum \( p \) and odd-sum \( q \), as explained above.

One can similarly establish formulae for product moments of more than two \( V(b)'s \), and investigate other aspects of the multivariate exponential distributions associated with
the $V(.)$ process. We remark that these distributions can be simulated, with some effort, through using $V_\varepsilon(b)$ with a small $\varepsilon$, and this is one way of computing bivariate and multivariate probabilities when needed. Another way would be via numerical inversion of the joint moment generating function.

4. Total relative time in $[a, \infty)$. As a generalisation of (1.1), consider the total relative time spent along the ray $w = bs$ during $s \geq a$, i.e.

$$V_\varepsilon(a, b) = \frac{1}{\varepsilon} \int_a^\infty I\{bs - \frac{1}{2}\varepsilon \leq W(s) \leq bs + \frac{1}{2}\varepsilon\} \, ds.$$  \hspace{1cm} (4.1)

The story told in the final paragraph of Section 1 is one motivation for studying these variables. The main result about them is that

$$V_\varepsilon(a, b) \rightarrow_d V(a, b) \sim k(|b|\sqrt{a})\exp(|b|) + (1 - k(|b|\sqrt{a}))\delta_0,$$  \hspace{1cm} (4.2)

where $\delta_0$ denotes unit point mass at zero and $k(c) = 2(1 - \Phi(c))$. Note that $k(|b|\sqrt{a}) = 1$ when $a = 0$, so that (4.2) indeed contains our earlier result (2.1).

To prove this, consider moments again. Take $b > 0$ for simplicity. It takes one moment to show that

$$EV_\varepsilon(a, b) \rightarrow \int_a^\infty f_s(bs) \, ds = \int_a^\infty \phi(b\sqrt{s})/\sqrt{s} \, ds = \frac{1}{b} \int_{b\sqrt{a}}^\infty 2\phi(x) \, dx = k(b\sqrt{a})/b.$$  \hspace{1cm} (4.3)

And when $p \geq 2$ we find

$$EV_\varepsilon(a, b)^p \rightarrow p! \int_a^\infty \cdots \int_{a \leq u_1 < \cdots < u_p} f_{s_1, \ldots, s_p}(bs_1, \ldots, bs_p) \, ds_1 \cdots ds_p$$
$$= p! \int_a^\infty \int_0^\infty \cdots \int_0^\infty \frac{\phi(b\sqrt{u_1})}{\sqrt{u_1}} \cdots \frac{\phi(b\sqrt{u_p})}{\sqrt{u_p}} \, du_1 \cdots du_p$$
$$= p! \frac{k(b\sqrt{a})}{b} \left(\frac{1}{b}\right)^{p-1}.$$  \hspace{1cm} (4.4)

The moment generating function of this limit distribution candidate becomes

$$E \exp(tV) = 1 + \sum_{p=1}^\infty \frac{t^p}{p!} p! k(b\sqrt{a}) \left(\frac{1}{b}\right)^{p-1}$$
$$= 1 + k(b\sqrt{a}) \frac{t/b}{1 - t/b} = k(b\sqrt{a}) \frac{1}{1 - t/b} + 1 - k(b\sqrt{a}),$$

which is recognised as the moment generating function of the mixture variable that with probability $k(b\sqrt{a})$ is an exponential with parameter $b$ and with probability $1 - k(b\sqrt{a})$ is equal to zero. This proves (4.2). $\square$

**Remark.** Let us briefly discuss a specific consequence, namely that $\Pr\{V_\varepsilon(a, b) = 0\}$ in this situation converges to $\Pr\{V(a, b) = 0\}$, which is $1 - k(b\sqrt{a}) = 2\Phi(b\sqrt{a}) - 1$. But having $V_\varepsilon(a, b) = 0$ in the limit means that $W(s)$ stays away from $bs$ during $[a, \infty)$, and it
cannot stay above the curve all the time since \( W(s)/s \) goes to zero. Hence \( 2\Phi(b\sqrt{a}) - 1 \) is simply the probability that \( W(s) < bs \) during all of \([a,\infty)\), or \( \Pr\{\max_{s \geq a} W(s)/s < b\} \). Using finally the transformation \( W^*(t) = tW(1/t) \) to another Brownian motion one sees that

\[
\Pr\{\max_{0 \leq t \leq 1/a} W(t) \leq b\} = 2\Phi(b\sqrt{a}) - 1 = \Pr\{|W(1/a)| \leq b\}.
\] (4.4)

We have in other words rederived a classic distributional result for Brownian motion. □

The distribution of \( V(a,-1) - V(a,1) \) comes up in the statistical estimation problem discussed in Section 1; see also 5C below and Hjort and Fenstad (1991b, Section 6). When \( a = 0 \) this is a difference between two unit exponentials with intercorrelation \(-\frac{1}{2}\). The case \( a > 0 \) is more complicated. Then

\[
(V(a,-1), V(a,1)) = \begin{cases} 
0 & \text{with probability } \pi_{00}, \\
(U_{-1}, 0) & \text{with probability } \pi_{10}, \\
(0, U_1) & \text{with probability } \pi_{01}, \\
(U_{-1}, U_1) & \text{with probability } \pi_{11},
\end{cases}
\] (4.5)

in which \( U_{-1} \) and \( U_1 \) are unit exponentials with a certain dependence structure. Furthermore \( \pi_{00} \) is the probability that \( W(s) \) stays between \(-s\) and \( s \) during \([a,\infty)\), \( \pi_{10} \) is the probability that \( W(s) \) comes below \(-s\) but is never above \( s \), \( \pi_{01} \) is the probability that \( W(s) \) comes above \( s \) but is never below \(-s\), and \( \pi_{11} \) is the probability that \( W(s) \) experiences both \( W(s) < -s \) and \( W(s) > s \) during \([a,\infty)\). When \( a = 0 \) then \( \pi_{11} = 1 \) and the others zero. In the positive case these probabilities can be found in terms of \( H(u) \), the probability that \( \max_{0 < s < 1} |W(s)| \leq u \), by the transformation arguments used to reach (4.4). One finds

\[
\pi_{00} = H(\sqrt{a}), \quad \pi_{01} = \pi_{10} = 2\Phi(\sqrt{a}) - 1 - \pi_{00}, \quad \pi_{11} = 1 - \pi_{00} - \pi_{01} - \pi_{10},
\]

in which \( H(u) = \Pr\{\max_{0 < s < 1} |W(s)| \leq u\} \). A classic alternating series expression for \( H(u) \) can be found in Shorack and Wellner (1986, Chapter 2, Section 2), for example, and a new way of deriving this formula is by calculating all product moments \( EV(-a)^p V(a)^q \) and then study the analogue of (3.6). This would be analogous to the way in which (4.4) was proved above, but the present case is much more laborious. Here we merely note that

\[
EV(a,-1) V(a,1) = \pi_{11} EU_{-1} U_1 = \frac{2}{3} k(3\sqrt{a}),
\]

from which the correlation between \( U(-1) \) and \( U(1) \) also can be read off.

5. Supplementing results.

5A. Total relative time along the time axis. The variable \( V_\varepsilon(b) \) of (1.1) is infinite when \( b = 0 \). But consider

\[
V_{\varepsilon,T} = \frac{1}{\varepsilon} \frac{1}{\sqrt{T}} \int_0^T \{-\frac{1}{2} \varepsilon \leq W(s) \leq \frac{1}{2} \varepsilon\} ds,
\] (5.1)
the relative time along the time axis during \([0,T]\). The moment sequence converges as 
\(\epsilon \to 0\) and \(T \to \infty\), as follows, using (2.3) once more:

\[
E(V_{\epsilon,T}^p) = \frac{p!}{\epsilon^p T^p/2} \int \cdots \int_{0<s_1<\cdots<s_p<T} \left[ \phi(0)^p \frac{\epsilon}{\sqrt{s_1}} \cdots \frac{\epsilon}{\sqrt{s_p - s_{p-1}}} + O(\epsilon^{p+1}) \right] ds_1 \cdots ds_p \\
\to p! \phi(0)^p \int \cdots \int_{0<s_1<\cdots<s_p<1} x_1^{-1/2} \cdots (x_p - x_{p-1})^{-1/2} (1 - x_p)^0 \, dx_1 \cdots dx_p \\
= \frac{p!}{(2\pi)^{p/2}} \frac{\Gamma(1/2)^p \Gamma(1)}{\Gamma(p/2 + 1)} = \frac{1}{2^{p/2}} \frac{p!}{\Gamma(p/2 + 1)}.
\]

The limit distribution candidate \(V_0\) has consequently \(EV_0^{2p} = (\frac{1}{2})^p (2p)! / p!\), which means that \(V_0^2\) gets moment generating function \((1 - 2t)^{-1/2}\). So \(V_0^2\) is a \(\chi^2_1\) (since the distribution of a chi-squared is determined by its moments), i.e. \(V_0\) is a \(|N(0,1)|\).

It wasn’t necessary here to send \(T\) to infinity, since the scaling property for \(W\)
\([W^*(s) = W(cs)/\sqrt{c}\) gives a new Brownian motion\] implies that the limit distribution of \(V_{\epsilon,T}\) is independent of \(T\).

\[5B. \text{Implications for partial sum processes.}\]

Let us first point out that an alternative construction of our total relative time variables is to use \(I\{bs \leq W(s) \leq bs + \epsilon\}\) instead of \(I\{W(s) \in bs \pm \frac{1}{2}\epsilon\}\) in (1.1) and (4.1). Results of previous sections hold equally for this alternative definition of \(V_{\epsilon}(b)\) and \(V_{\epsilon}(a,b)\), and this is a bit more convenient in 5C below.

Now suppose \(X_1, X_2, \ldots\) are i.i.d. with mean \(\xi\) and variance \(\sigma^2\), and consider the normalised partial sum process \(W_m(s) = m^{-1/2} \sum_{i=1}^{[m]} (X_i - \xi) / \sigma\). In particular \(W_m(n/m) = S_n / \sqrt{m}\), writing \(Y_i = (X_i - \xi) / \sigma\) and \(S_n\) for their partial sums, and \(W_m(.)\) converges to Brownian motion by Donsker’s theorem. Motivated by (1.1) and (4.1) we define

\[
V_{m,\epsilon}(a, b) = \frac{1}{\epsilon} \sum_{n/m \geq a} I\left\{ \frac{n}{m} \leq \frac{S_n}{\sqrt{m}} \leq b \frac{n}{m} + \epsilon \right\} \\
= \frac{1}{\epsilon} \sum_{n/m \geq a} I\left\{ b \sqrt{\frac{n}{m}} \leq T_n = \sqrt{n}(X_n - \xi) / \sigma \leq b \sqrt{\frac{n}{m}} + \epsilon \sqrt{\frac{m}{n}} \right\} \\
= \frac{1}{\epsilon} \int_{(am)/m}^{\infty} I\left\{ b \frac{[ms]}{m} \leq W_m(s) \leq b \frac{[ms]}{m} + \epsilon \right\} \, ds,
\]

where \(\langle am \rangle\) denotes the smallest integer exceeding or equal to \(am\). It is clear that this variable is close to \(V_{\epsilon}(a,b)\) for large \(m\), and should accordingly converge in distribution to \(V(a,b)\) of (4.2) when \(m \to \infty\) and \(\epsilon \to 0\).

\[\text{PROPOSITION. Assume that the } X_i\'s \text{ have a finite third absolute moment. If } a > 0 \text{ is fixed, then } V_{m,\epsilon}(a,b) \to_d V(a,b) \text{ if only } \epsilon(m) \to 0 \text{ as } m \to 0. \text{ And}
\]

\[
V_{m,\epsilon}(a(m), b) \to_d V(b) \sim \text{Exp}(|b|)
\]

provided \(\epsilon(m) \to 0, a(m) \to 0, ma(m) \to \infty, \text{ and } \epsilon(m)/a(m)^{1/2} \to 0.\)

This can be proved in various ways and under various conditions. One feasible possibility is to demonstrate moment convergence of \(E\{V_{m,\epsilon}(a,b)\}^p\) towards the right hand
side of (4.3), for each $p$. One basically needs the smallest $n$ in the sum to grow towards infinity, so that the central limit theorem and Edgeworth expansions can begin to work, and the largest of all $\epsilon(m)\sqrt{m/n}$ terms to go to zero, so that Taylor expansions can begin to work; see the middle expression (5.2). When $a$ is fixed then the sum is over all $n \geq ma$, and it suffices to have $\epsilon(m) \to 0$ as $m \to \infty$. To reach $V(b) = V(0,b)$ in the limit we need the stated behaviour for $\epsilon(m)$ and $a(m)$. I have used the third moment assumption to bound the error $r(t)$ in the Edgeworth expression $G_n(t) = \Phi(t) + r(t)$ for the distribution of $T_n = \sqrt{n}(\bar{X}_n - \xi)/\sigma$; one has $|r_n(t)| \leq cn^{-1/2}/(1+|t|^3)$, and this is helpful when it comes to verifying conditions when employing Lebesgue’s theorem on dominated convergence.

We may conclude that the total relative time along $b\xi$, for the normalised partial sum process has a limit distribution, which is either exponential or of the mixture type (4.2). The middle expression also invites $V_{m,\epsilon}$ to be thought of as the total relative time for the normalised $T_n$ process along the square root boundary $b\sqrt{n}/m$. The result is also valid for $T_n = \sqrt{n}(\theta_n - \theta)/\sigma$ in a more general estimation theory setup; see Hjort and Fenstad (1991a, 1991b).

The 5A result has also implications for partial sum processes. One can prove that
\[
\frac{1}{\epsilon} \int_0^1 I\{|W_m(s)| \leq \frac{1}{\epsilon}\} \, ds \to_d |N(0,1)|
\]
when $\epsilon \to 0$ and $m \to \infty$, under suitable conditions. This implies for example that $m^{-1/2} \sum_{i=1}^m I\{|S_i| \leq \frac{1}{2}\}$ has the absolute normal limit, as does $m^{-1/2} \sum_{i=1}^m I\{S_i = 0\}$ for the random walk process.

5C. Second order asymptotics for the number of $\delta$-errors. To show how the total relative time variables for Brownian motion are related to the estimation theory problems described in Section 1, consider the structurally simple case of i.i.d. variables $X_i$ with mean $\xi$ and standard deviation $\sigma$, and where $\frac{n}{n+c} \bar{X}_n$ is used to estimate $\xi$. Consider $Q_\delta(c)$, the number of times $|\frac{n}{n+c} \bar{X}_n - \xi| \geq \delta$, counted among $n \geq a/\delta^2$. Then $\delta^2 Q_\delta(c)$ tends to $Q = Q(a,1/\sigma)$ of (1.2), for each choice of $c$, and $\delta^2$ times $Q_\delta(c) - Q_\delta(0)$ goes to zero. This follows from results in Hjort and Fenstad (1991a). But $\delta (Q_\delta(c) - Q_\delta(0))$ can be written $A_\delta - B_\delta$, after some analysis, where
\[
A_\delta = \sqrt{m} \int_{(ma)/m}^{\infty} I\{\frac{[ms]}{m} \frac{1}{\sigma} \leq W_m(s) \leq \frac{[ms]}{m} \frac{1}{\sigma} + \frac{1}{\sqrt{m}} \frac{c\xi}{\sigma} - \frac{1}{m} \frac{c}{\sigma}\} \, ds,
\]
\[
B_\delta = \sqrt{m} \int_{(ma)/m}^{\infty} I\{\frac{[ms]}{m} \frac{1}{\sigma} \leq W_m(s) \leq \frac{[ms]}{m} \frac{1}{\sigma} + \frac{1}{\sqrt{m}} \frac{c\xi}{\sigma} + \frac{1}{m} \frac{c}{\sigma}\} \, ds,
\]
and where $m = 1/\delta^2$. These variables resemble those considered in (5.2) and (5.3). With $\epsilon = m^{-1/2}c\xi/\sigma$ we have
\[
A_\delta \overset{d}{=} c\xi/\sigma V_{m,\epsilon}(a,-1/\sigma) \quad \text{and} \quad B_\delta \overset{d}{=} c\xi/\sigma V_{m,\epsilon}(a,1/\sigma),
\]
where ‘$\overset{d}{=}$’ signifies that the difference goes to zero in probability. It follows from the result of 5B that
\[
\delta (Q_\delta(c) - Q_\delta(0)) \to_d c\xi/\sigma \{V(a,-1/\sigma) - V(a,1/\sigma)\} \quad \text{as } \delta \to 0. \quad (5.4)
\]
This is also true with \( a = 0 \) in the limit, i.e. with \( c \xi / \sigma \{ V(-1/\sigma) - V(1/\sigma) \} \) on the right hand side, provided \( a = a(\delta) = \delta \) is used in the definition of \( Q_\delta(c) \) and \( Q_\delta(0) \). Note the relevance of (4.5) for the present problem.

Hjort and Fenstad (1991b) also work with the direct expected value of \( Q_\delta(c) - Q_\delta(0) \) and similar variables. These converge to explicit functions of \( c \) (and other parameters, in more general situations), which can then be minimised to single out estimator sequences with the second order optimality property of having the smallest expected number of \( \delta \)-errors. This is done in Hjort and Fenstad (1991b), in several situations. We remark that the skewness \( \gamma = E(X_i - \xi)^3 / \sigma^3 \) is not involved in (5.4), but is prominently present in the limit of \( E\{Q_\delta(c) - Q_\delta(0)\} \), and its minimisation.

5D. Total relative time along other curves. A generalisation of (1.1) and (4.1) is to replace \( bs \) with a general function \( b(s) \), say \( bs^{1/2} \) or \( bs^2 \). The resulting \( V_\varepsilon \) would again converge in distribution to some appropriate \( V \), the total relative time Brownian motion spends along the curve \( w = b(s) \). One can typically compute the first couple of moments of \( V \), in the general case, but its distribution is simple only in the linear case. This gives in particular an (admittedly esoteric) characterisation of the exponential distribution.

5E. Other Markov processes. (i) If \( X(t) = \mu t + W(t) \) is Brownian motion with drift \( \mu \) we have immediate results for \( V^*(b) \) and \( V^*(a, b) \), the total relative time variables along \( x = bs \), since \( V^*(b) = V(b - \mu) \) and \( V^*(a, b) = V(a, b - \mu) \) from the definitions. (ii) To generalise our framework, consider

\[
V_\varepsilon = \frac{1}{\varepsilon} \int_0^\infty I\{b(s) \leq X(s) \leq b(s) + \varepsilon g(s)\} \, ds,
\]

where \( X(.) \) is a Markov process, \( x = b(s) \) some curve of interest, and \( g(s) \) a possible scaling factor. In many cases there is a distributional limit as \( \varepsilon \to 0 \), and perhaps the first couple of moments can be obtained. The limit distribution is simple only for cases that can be transformed back to (1.1) and (4.1), however, cf. 5D above. So the total relative time an Ornstein–Uhlenbeck process \( X(s) \) spends along \( be \) can be shown to be exponential, for example, with suitable \( g(s) \) in (5.5).

5F. Other exponential and gamma processes. (i) If \( U(b) = |b|V(b) \), then \( U(b) \) is unit exponential for each \( b \). In particular its marginal mean and variance are constant, and \( \text{cov}\{U(b), U(b+h)\} = b/(b+2h) \). (ii) By adding independent copies of \( V(.) \) (or \( U(.) \)) we get processes with marginals that are gamma distributed. This leads in particular to bivariate and multivariate gamma distributions or chi-squared distributions (with even-numbered degrees of freedom only). (iii) There are other processes that share with \( V \) and \( U \) the property of having exponentially distributed marginals. An example is \( V^*(b) = \frac{1}{2}\{W_1(b)^2 + W_2(b)^2\} \), where \( W_1 \) and \( W_2 \) are independent Brownian motions. This is a Markov process, while our \( V(b) \) process is not. The possible correlations of \((V^*(b_1), \ldots, V^*(b_n))\) span a smaller space than those of \((V(b_1), \ldots, V(b_n))\), indicating that the \( V^* \) process may be less adequate when it comes to building multivariate exponential models.
References

Hjort, N.L. and Fenstad, G. (1991a). On the last time and the number of times an estimator is more than $\epsilon$ from its target value. To appear in *Annals of Statistics*.

Hjort, N.L. and Fenstad, G. (1991b). Some second order asymptotics for the number of times an estimator is more than $\epsilon$ from its target value. Statistical research report, University of Oslo; submitted for publication.