MA 8101

Comments on Girsanov's Theorem

Harald E. Krogstad, IMF 2003

Girsanov's Theorems (presented in three different versions in B.Ø.) are probability measure transformations applied to Ito processes. These transformations do not change the solutions, but since the underlying probability measures will be different, the defining equations will have different drift coefficients and new versions of the Brownian motion. The objective of the transformation will typically be to obtain more favorable probability measures, e.g. in order to ease simulation studies.

Assume that we have an Ito model of the form

$$dX_t(\omega) = b(t,\omega)\,dt + \sigma(t,\omega)\,dB_t. \tag{1}$$

We want to find the probability that X_t reaches some critical domain D during a time interval $[0,T]$, and we know that this probability is quite small. The only way for us is to carry out numerical simulations, but if the probability is, say 10^{-6}, we would have to carry out more than 10^6 simulations in order to get a reasonable estimate for it. We could change the drift $b(t,\omega)$ in Eqn. 1 to, say $a(t,\omega)$, so that for the modified process,

$$dY_t(\omega) = a(t,\omega)\,dt + \sigma(t,\omega)\,dB_t, \tag{2}$$

the event that Y_t enters D during $[0,T]$ is quite likely. To estimate the corresponding probability by simulations using Y_t is therefore simple, requiring only a few realizations. So far this is not surprising, but Girsanov's Theorem tells us that it may, in lucky cases, be possible to change the underlying probability measure so that the solutions of Eqn. 2 also satisfy a model

$$d\tilde{X}_t(\omega) = b(t,\omega)\,dt + \sigma(t,\omega)\,d\tilde{B}_t, \tag{3}$$

similar to Eqn. 1. Here $\tilde{B}_t$ is a Brownian motion with respect to the new probability measure. Assuming weak uniqueness for Eqn. 1, the probability laws for the solutions of Eqns. 1 and 3 will thus be the same. In particular, the probability we are looking for would be the same for 1 and 3.

Girsanov's Theorem provides an explicit expression for the measure transformation, and we may in fact carry out the simulations by means of 2 (where the event is quite likely) and then compute the estimate of the probability we are seeking for model 1 (or 3) by means of the outcome from the simulations and the expression for the measure transformation. This will be described in more details below, and the idea is related to what is called *importance sampling* in statistics.

1 The Radon-Nikodym Theorem

Let $\{\Omega, \mathcal{F}, P\}$ be a probability space and M a non-negative $\mathcal{F}$-measurable stochastic variable such that $\int_\Omega M(\omega)\,dP(\omega) = 1$. We may then define a new probability measure Q on Ω by

$$dQ(\omega) = M(\omega)\,dP(\omega), \tag{4}$$

that is,

$$\int_\Omega h(\omega)\,dQ(\omega) = \int_\Omega h(\omega)\,M(\omega)\,dP(\omega) \tag{5}$$

for all $\mathcal{F}$-measurable functions h such that the integrals exist. In this case, $\{\Omega, \mathcal{F}, Q\}$ is also a probability space, but in general, $Q(A) \neq P(A)$ and $L^p(\Omega, \mathcal{F}, Q) \neq L^p(\Omega, \mathcal{F}, P)$. Moreover, if $X : \Omega \to \mathbb{R}^n$ is a stochastic variable, then the distributions of X with respect to P and Q will be different.

Let P and Q be two measures on Ω. If $P(A) = 0 \implies Q(A) = 0$ for all $A \in \mathcal{F}$ with zero P-measure, then we say that Q is *absolutely continuous* with respect to P and write this as $Q \ll P$.

In the case above, if $P(A) = 0$, then

$$Q(A) = \int \chi_A(\omega) M(\omega) dP(\omega) = 0, \tag{6}$$

since $\chi_A(\omega) M(\omega) = 0$ a.s.(P). Hence, $Q \ll P$ in this case. The opposite will not be true in general: If $M(\omega) = 0$ on a set A where $P(A) > 0$, then $Q(A) = \int_A M(\omega) dP(\omega) = 0$, but this does not imply that $P(A) = 0$. However, if $M(\omega)$ is strictly positive for all ω, then it works both ways, and $Q \ll P$ and $P \ll Q$.

The *Radon-Nikodym Theorem* is a striking converse to the above situation:

If $Q \ll P$, there exists a unique a.e. (P) $\mathcal{F}$-measurable function M such that Eqn. 4 holds.

It is customary to write

$$M = \frac{dQ}{dP}, \tag{7}$$

and call M the *Radon-Nikodym derivative.*

2 Importance sampling

Assume that X is a stochastic variable in $\mathbb{R}^n$ on the probability space $(\Omega, \mathcal{F}, P)$ with values in $\mathbb{R}^n$, and with a probability distribution $\mu_X^{(P)}$. The expectation of $h(X)$ with respect to P is (when it exists)

$$m^{(P)} = E^{(P)}(h(X)) = \int_\Omega h(X(\omega)) dP(\omega) = \int_{\mathbb{R}^n} h(x) d\mu_X^{(P)}(x). \tag{8}$$

In a practical situation, we may want to *estimate* this expectation from a series of observations or simulations of X, say $\{x_i\}_{i=1}^S$. The obvious estimate is

$$\hat{m}^{(P)} = \frac{1}{S}\sum_{i=1}^{S} h(x_i), \tag{9}$$

where the observations are supposed to be drawn from the distribution of X under P.

We could also consider the same problem under a transformed measure Q, and let us assume that Q is absolutely continuous with respect to P so that

$$dQ(\omega) = M(\omega) dP(\omega). \tag{10}$$

We are going to assume that the observations that gives us x_i also provide observations of M at the same time, say m_i. As an example, assume that that X under P and Q have strictly positive probability densities,

$$\begin{aligned} d\mu_X^{(P)}(x) &= f^{(P)}(x) dx, \\ d\mu_X^{(Q)}(x) &= f^{(Q)}(x) dx. \end{aligned} \tag{11}$$

Then

$$M(\omega) = \frac{f^{(Q)}(X(\omega))}{f^{(P)}(X(\omega))} \tag{12}$$

defines an acceptable transformation, since $M(\omega) > 0$ and

$$E^{(P)}(M) = \int_{\mathbb{R}^n} \frac{f^{(Q)}(x)}{f^{(P)}(x)} f^{(P)}(x)\, dx = \int_{\mathbb{R}^n} f^{(Q)}(x)\, dx = 1. \tag{13}$$

Assume that we want an estimate of the Q-probability that X falls in a set D, that is, the Q-measure of the set $\{\omega\ ;\ X(\omega) \in D\}$. This may be written

$$Q\{\omega\ ;\ X(\omega) \in D\} = \int_{X(\omega)\in D} dQ(\omega) = \int_{\Omega} \chi_D(X(\omega))\, dQ(\omega) = E^{(Q)}(\chi_D(X)). \tag{14}$$

If the probability is very small, one would need a correspondingly large number of simulations in order to find a reasonable estimate of the probability. However, assume that the corresponding probability of X falling in D when X is drawn according to $d\mu_X^{(P)}$ is about $1/2$. Then the event $X(\omega) \in D$ would occur for about half of the simulations, and we could do with a relatively small number of observations for obtaining an accurate estimate of $P(\omega\ ;\ X(\omega) \in D)$. Assume, as discussed above, that we obtain M along with X (drawn according to P). We then observe that

$$\begin{aligned} E^{(Q)}(\chi_D(X)) &= \int_{\Omega} \chi_D(X(\omega))\, dQ(\omega) \\ &= \int_{\Omega} \chi_D(X(\omega))\, M(\omega)\, dP(\omega) \\ &= E^{(P)}(\chi_D(X)\, M). \end{aligned} \tag{15}$$

It is therefore also possible to estimate $E^{(Q)}(\chi_D(X))$ from observations of X and M according to P. This suggests the alternative estimate

$$\hat{m}^{(Q)} = \frac{1}{S} \sum_{i=1}^{S} \chi_D(x_i)\, m_i. \tag{16}$$

Since $\chi_D(x_i)$ now will be non-zero for a considerable fraction of the observations, and M is in general a non-zero function, the estimate will quickly approach a reliable value when S increases. In the example above, where M is given in Eqn. 12, the formula amounts to

$$\hat{m}^{(Q)} = \frac{1}{S} \sum_{i=1}^{S} \chi_D(x_i) \frac{f^{(Q)}(x_i)}{f^{(P)}(x_i)}, \tag{17}$$

and this is denoted *importance sampling* is statistics.

3 Girsanov's Theorems

The theorem is stated in various forms in B.Ø., but before we start it is recommended to take a look at Exercise 4.4 about *Exponential Martingales*. The exercise discusses stochastic process of the form

$$Z_t = \exp\left\{ \int_0^t \theta(s,\omega)' \cdot dB_s - \frac{1}{2} \int_0^t |\theta(s,\omega)|^2\, ds \right\},\ 0 \le t \le T, \tag{18}$$

where B_s and θ are in $\mathbb{R}^n$, and $\theta \in \mathcal{W}^n(0,T)$, so that all integrals are defined. These processes are Martingales with respect to the filtration of the Brownian motion and P under some rather weak additional conditions on θ. One such condition is the *Novikov Condition*:

$$E\left(\exp\left(\int_0^T \frac{1}{2}|\theta(s,\omega)|^2 ds\right)\right) < \infty. \tag{19}$$

(See Exercise 4.4 for references).

The first version of Girsanov's Theorem in B.Ø. concerns the rather simple case where the Ito process has the form

$$dY_t(\omega) = a(t,\omega)\,dt + dB_t(\omega) \tag{20}$$

and where a is adapted to the Brownian motion. Define

$$M_t = \exp\left\{-\int_0^t a(s,\omega)' \cdot dB_s - \frac{1}{2}\int_0^t |a(s,\omega)|^2 ds\right\},\ 0 \le t \le T, \tag{21}$$

and assume that M_t is a Martingale (E.g. a satisfies the Novikov Condition). Let

$$dQ = M_T dP, \tag{22}$$

defined on $\mathcal{F}_T$. Then Q is a probability measure since it is clearly non-negative and

$$Q(\Omega) = \int_\Omega M_T dP = \int_\Omega E(M_T|\mathcal{M}_0)\,dP = \int_\Omega M_0 dP = 1. \tag{23}$$

Moreover, it is worth noting that for all $\mathcal{F}_t$-measurable functions, $t < T$,

$$dQ = M_t dP, \tag{24}$$

because of the Martingale property of M_t (Exercise for the reader, or see B.Ø.).

The Ginsarov's theorem now states that *the solution of the process in Eqn. 20 is a Brownian motion with respect to the modified measure Q defined in* 22.

The proof consists of verifying that Y_t satisfies the Lévy conditions stated in B.Ø., Theorem 8.6.1 with respect to Q.

We see that even if Y_t has a drift, M_t shifts the weight of the probability so that it "follows" Y_t, and Y_t becomes a (standard) Brownian motion with respect to the new measure Q.

3.1 A worked example

The idea of this example is taken from [2], pp. 337-340.

In order to see how Girsanov's Theorem works, consider the following simple example:

$$dY_t = dt + dB_t,\ Y_0 = x. \tag{25}$$

The corresponding Martingale M_t is, according to Eqn. 21,

$$M_t = \exp\left(-\int_0^t dB_s - \frac{1}{2}\int_0^t ds\right) = \exp\left(-B_t - \frac{t}{2}\right), \tag{26}$$

and

$$dQ(\omega) = \exp\left(-B_T(\omega) - \frac{T}{2}\right) dP(\omega). \tag{27}$$

Since the solution of 25 is

$$Y_t = x + t + B_t, \tag{28}$$

we have

$$E^{(P)} Y_t = x + t \tag{29}$$

On the other hand, since Y_t is a Brownian motion with respect to Q,

$$E^{(Q)} Y_t = x. \tag{30}$$

Let us verify 30 by using the expression for Q in Eqn. 27:

$$\begin{aligned}
E^{(Q)} Y_t &= \int_{\Omega} Y_t dQ \\
&= \int_{Q} (x + t + B_t) \exp\left(-B_t - \frac{t}{2}\right) dP(\omega) \\
&= \int_{\mathbb{R}} (x + t + \eta) \exp\left(-\eta - \frac{T}{2}\right) \frac{1}{\sqrt{2\pi t}} e^{-\frac{\eta^2}{2t}} d\eta \\
&= \int_{\mathbb{R}} (x + t + \eta) \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(\eta + t)^2}{2t}\right) d\eta \\
&= x + t + \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} \eta \exp\left(-\frac{(\eta + t)^2}{2t}\right) d\eta \\
&= x + t + \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} (\eta + t - t) \exp\left(-\frac{(\eta + t)^2}{2t}\right) d\eta \\
&= x + t - t = x.
\end{aligned} \tag{31}$$

(Check that we could as well have used $dQ = M_T dP$ instead of $dQ = M_t dP$!).

This case result was trivial, so let us consider the more challenging problem of finding the probability that a standard Brownian motion starting at 0 exceeds a level K in the interval $[0, 1]$.

Since the process Y_t for $Y_0 = 0$ is a standard Brownian with respect to Q, the probability we are looking for is

$$p = Q\left\{\omega\ ;\ \sup_{0 \leq t \leq 1} Y_t(\omega) > K\right\}.$$

In this special case, there exists an analytic solution, but it is obvious:

$$p = \sqrt{\frac{2}{\pi}} \int_K^{\infty} e^{-x^2/2} dx \tag{32}$$

(see [1], Sec. 26). As pointed out during the lecture by Sergei, this results may be obtained (or rather made plausible) by using a so-called *reflection principle*. The principle works excellent for a discrete random walk, where the number of possible paths is finite. It is not so obvious to apply for a case where the number of paths is not even countable. Nevertheless, the same result is true, but the rigorous mathematical proof is not simple. If we do not happen to know this result, the only way for us to find p is to use *simulations*. However, if $K \gg 1$, the probability p is *very* small, and we would need a lot of simulations in order to determine its value. On the other hand, the process Y_t under P has a *drift*, and we may adjust this drift so that a reasonable fraction of the paths simulated from Eqn. 25 exceeds K. Let us therefore consider the slightly modified process

$$dY_t = Kdt + dB_t,\ Y_t = 0,\ 0 \leq t \leq 1, \tag{33}$$

so that

$$Y_t = Kt + B_t. \tag{34}$$

Since $E^{(P)}(Y_1) = K$, we expect that more than half of the paths have a maximum above K (We would, however, like to avoid a situation where virtually all paths exceed K). The Radon-Nikodym derivative is

$$\frac{dQ}{dP} = M_1 = \exp\left(-KB_1 - \frac{K^2}{2}\right), \tag{35}$$

and as soon as we have simulated the Brownian motion and found a sample of Y_t from Eqn. 34, we also have M_1 for the same sample.

Below we shall set K to different values and carry out computer experiments using a discrete Brownian motion with $\Delta t = 10^{-5}$. Let $N(\omega)$ be the random function

$$N(\omega) = \begin{cases} 1, & \max_{0\leq t\leq 1} Y_t(\omega) \geq K, \\ 0, & \max_{0\leq t\leq 1} Y_t(\omega) < K. \end{cases} \tag{36}$$

We are looking for $E^{(Q)}(N)$, and there are two ways of proceeding. The first is to simulate Y_t under Q, that is, a standard Brownian motion. We carry out S simulations, note the number of times the simulations exceed K, say s times, and set

$$\hat{p}_1 = \frac{s}{S}.$$

The second way is to use the same simulated Brownian motion as input for simulated solutions $y_i(t)$ of Eqn. 34. Let

$$\begin{aligned} n_i &= \begin{cases} 1, & \max_{0\leq t\leq 1} y_i(t) \geq K, \\ 0, & \text{otherwise}, \end{cases} \\ b_i &= B_i(1), \end{aligned} \tag{37}$$

$i = 1, \cdots, S$. From the formula

$$\begin{aligned} E^{(Q)}(N) &= \int_\Omega N(\omega)\, dQ(\omega) \\ &= \int_\Omega N(\omega)\, M_1(\omega)\, dP(\omega) \\ &= E^{(P)}(NM_1), \end{aligned} \tag{38}$$

we obtain the alternative estimate based on Eqns. 34 and 35,

$$\hat{p}_2 = \frac{1}{S}\sum_{i=1}^{S} n_i \exp\left(-Kb_i - \frac{K^2}{2}\right). \tag{39}$$

In this case we expect that $n_i = 1$ for more than half of the simulations. Table 1 shows the outcome when we have used 500 simulations. For $K = 1$ and 2, both estimates are

K	p_{exact}	$\hat{p}_1$	$\hat{p}_2$	$\hat{p}^{(P)}$
1	0.32	0.31	0.33	0.67
2	0.046	0.034	0.042	0.62
3	2.7×10^{-3}	2×10^{-3}	2.9×10^{-3}	0.61
4	6.3×10^{-5}	0	5.8×10^{-5}	0.57
5	5.7×10^{-7}	0	5.2×10^{-7}	0.54

Table 1: Results of 500 computer sumulations for various values of K. The exact values of the probability is computed by Eqn. 32, whereas $\hat{p}_1$ and $\hat{p}_2$ are defined in the text. The rightmost column shows the probability of exceedence under P.

reasonable. However, already for $K = 3$, the direct estimate shows that only one sample out of 500 exceeded the limit. For $K = 4$ and 5, none of the direct simulations exceeded the limit, whereas about 55% of the modified simulations using Eqn. 34 did that. The alternative estimate $\hat{p}_2$ continues to work reasonably well, whereas $\hat{p}_1$ is useless. Note that p-values are subject to random sampling errors, and, moreover, our way of simulating the Brownian motion could introduce some bias in the estimates.

3.2 A more general version

Theorem 8.6.6 in B.Ø. presents a more general version of Girsanov's Theorem (The theorem reduces to the version above as a special case). The starting point is an Ito diffusion of the form

$$dX_t = b\left(X_t\right)dt + \sigma\left(X_t\right)dB_t. \tag{40}$$

It is possible to introduce an additive change in the drift $b\left(X_t\right)$,

$$\gamma_t + b\left(X_t\right), \tag{41}$$

provided that the process γ_t is adapted (to B_t) and there is an adapted solution u_t to the equation

$$\sigma\left(X_t\right)u_t = \gamma_t. \tag{42}$$

The solution of the modified process

$$dX_t = \left(\gamma_t + b\left(X_t\right)\right)dt + \sigma\left(X_t\right)dB_t \tag{43}$$

under P will also satisfy the equation

$$d\tilde{X}_t = b\left(\tilde{X}_t\right)dt + \sigma\left(\tilde{X}_t\right)d\tilde{B}_t, \tag{44}$$

with respect to the modified probability measure

$$\begin{aligned} dQ &= M_T dP, \\ M_t &= \exp\left\{-\int_0^t u_s' \cdot dB_s - \frac{1}{2}\int_0^t \left|u_s\right|^2 ds\right\}, \end{aligned} \tag{45}$$

and the Brownian motion

$$\tilde{B}_t = \int_0^t u_s ds + B_t,\ 0 \le t \le T. \tag{46}$$

Again, assuming weak uniqueness, the solutions of the process in Eqn. 44 has the same distributions as the process in Eqn. 40.

References

[1] Lamperti, J.: *Probability, a survey of the mathematical theory*, Benjamin, 1966.

[2] Grigoriu, M.: *Stochastic calculus: applications in science and engineering*, Birkhäuser, 2002.