Abstract

The question of this paper is simple: why do we want to change Measure and how it can be done?

To briefly answer the first question consider standard Black-Scholes model.

$$dS_t = \mu_t S_t \, dt + \sigma_t S_t \, dW_t \quad (\Omega, \mathcal{F}, P)$$

One of the reasons why Black, Scholes and Merton where awarded Nobel prize in Economics is that because they provided formula which showed how derivative can be hedged. In other words, they demonstrated how by trading in underlying and risk-free asset one can replicate the payoff of derivative. For now, assume $r = 0$.

$$V_T = V_0 + \int_0^T H_t dS_t$$

.
where V_0 is the price of derivative set at $t = 0$. But how can it be determined? One idea is to take expectation on both sides of above but the problem is we don't know μ in dS_t.
Now for a second think that we constructed probability measure Q such that discounted stock price $e^{-rt}S_t$ is a martingale. For case $r = 0$, this means that $E^Q[dS_t] = 0$. Then in this measure Q called risk-neutral

$$E^Q[V_T] = E^Q[V_0 + E\int_0^T H_t dS_t] = E^Q[V_0] + E^Q[\int_0^T H_t dS_t] = V_0$$

In other words, we have managed to determine the price of derivative without knowing μ. But how to construct and prove the existence of this measure Q?

We will try to answer these question by first examining Continuous time model and then referring to it's discretized version. (I believe it will be more clear in this order). After it we will refer to some of it's applications and provide examples.

Theoretical minimum

Throughout this paper we will heavily rely on theoretical concepts and results. Therefore, we briefly state some of them below.

Moment-Generating function

Let's consider Taylor series for random variable X and dummy variable t:

$$e^{tX} = 1 + t \cdot X + \frac{1}{2} \cdot t^2 X^2 + \frac{1}{6} \cdot t^3 X^3 + ... = \sum_{n=0}^{\infty} \frac{(tX)^n}{n!}$$

By taking expectation on the right and left-handside we obtain:

$$\mathrm{E}[e^{tX}] = 1 + t \cdot \mu_1 + \frac{1}{2} \cdot t^2 \mu_2 + \frac{1}{6} \cdot t^3 \mu_3 + ... = \sum_{n=0}^{\infty} \frac{(t)^n \cdot \mu_n}{n!}$$

where

$$\mu_n = \mathrm{E}[X^n]$$

The idea of moment generating function is that μ_n it completely determines the distribution of random variable (i.e. there is one-to-one correspondence between distribution and MGF).
This can be summarized by the following:

$$M_t(X) = M_t(Y) \leftrightarrow X \overset{d}{=} Y$$

Example
Let X $\sim N(\mu, \sigma^2)$. Then $M_t(X) = e^{ut + \frac{\sigma^2}{2} t^2}$

$M_t(X) = \int_{-\infty}^{\infty} e^{tx} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = e^{t\mu} \int_{-\infty}^{\infty} e^{t(x-\mu)} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$

$= e^{t\mu} \int_{-\infty}^{\infty} e^{t(x-\mu) + \frac{\sigma^2}{2} t^2 - \frac{\sigma^2}{2} t^2} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$

$= e^{t\mu + \frac{\sigma^2}{2} t^2} \int_{-\infty}^{\infty} e^{ta - \frac{\sigma^2}{2} t^2} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{a^2}{2\sigma^2}} da$

$= e^{t\mu + \frac{\sigma^2}{2} t^2} \int_{-\infty}^{\infty} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(a - t\sigma^2)^2}{2\sigma^2}} da = e^{t\mu + \frac{\sigma^2}{2} t^2} \int_{-\infty}^{\infty} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z)^2}{2\sigma^2}} dz$

$= e^{t\mu + \frac{\sigma^2}{2} t^2}$

Brownian motion
Brownian motion $W(t)_{t\geq 0}$ is a stochastic process which satisfies the following properties:

1. W(0)=0
2. $W(t) - W(s)$ is independent of filtration $\mathcal{F}_\mathrm{s}$ for any $s \leq t$
3. $W(t) - W(s) \sim N(0, t-s)$
4. $W(t)$ has continuous paths

Quadratic Variation
One prominent property of Brownian motion is that its quadratic variation is not equal to zero. This fact can be captured by:

$$dW_t \cdot dW_t = dt$$

Proof (not strict)
Quadratic Variation from 0 to T

$[W,W]_0^T = \lim_{n\to\infty} \sum_{i=1}^n \left(W\left(\frac{iT}{n}\right) - W\left(\frac{(i-1)T}{n}\right)\right)^2 = \lim_{n\to\infty} \sum_{i=1}^n \frac{n}{n} \left(W\left(\frac{iT}{n}\right) - W\left(\frac{(i-1)T}{n}\right)\right)^2$

$= \lim_{n\to\infty} n \cdot \sum_{i=1}^n \frac{\left(W\left(\frac{iT}{n}\right) - W\left(\frac{(i-1)T}{n}\right)\right)^2}{n} \overset{\text{LLN}}{=} \lim_{n\to\infty} n \cdot \frac{T}{n} = T$
$[W,W]_0^T = \lim_{n\to\infty} \sum_{i=1}^n \left(W\left(\frac{iT}{n}\right) - W\left(\frac{(i-1)T}{n}\right)\right)^2 = \int_0^T (dW_t)^2 = T$

Theorem below will be used later to show that under changed measure new process is indeed Brownian motion

Levy theorem, one dimension
Let $M(t)$ be:

1. martingale with respect to filtration $\mathcal{F}_\mathrm{s}$
   M(0)=0
2. $M(t)$ has continuous paths
3. $[W,W]_0^T$=T

Then $M(t)$ is a Brownian motion

Proof
Let's apply Ito's formula for $f(t, x = M(t))$. Then:

$$df(t, M(t)) = f_t(t, M(t))dt + f_x(t, M(t))dM_t + \frac{1}{2} f_{xx}(t, M(t))dt$$

where in the last term $dM_t \cdot dM_t = dt$ which accounts for quadratic variation.
In integral form

$$f(t, M(t)) = f(0, M(0)) + \int_0^t \left(f_t(s, M(s)) + \frac{1}{2} f_{xx}(s, M(s)) \right) ds + \int_0^t f_x(s, M(s))dM_s$$

Since M_t is a martingale the last integral is also a martingale. Therefore by taking expectation we get:

$$\mathrm{E}(f(t, M(t)) = f(0, M(0)) + \mathrm{E} \int_0^t \left(f_t(s, M(s)) + \frac{1}{2} f_{xx}(s, M(s)) \right) ds$$

Now let's consider function

$$f(t, x) = e^{ux - \frac{1}{2}u^2 t}$$

For it

$$f_t(s, M(s)) + \frac{1}{2} f_{xx}(s, M(s)) = 0$$

Therefore

$$\mathrm{E}(f(t, M(t)) = \mathrm{E}\left(e^{uM(t) - \frac{1}{2}u^2 t} \right) = f(0, M(0)) = 1$$

or

$$\mathrm{E}\left(e^{uM(t)} \right) = e^{\frac{1}{2}u^2 t}$$

which coincides with MGF for Normal distribution with mean zero and variance of t. Therefore $M(t)$ should have the same distribution as $N(0, t)$.

Radon-Nikodym derivative

Properties of random variable depend on the probability which we assign to different outcomes. Radon-Nikodym derivative provides theoretical basis for this.

Radon-Nikodym derivative
Let X be a random variable under triple $(\Omega, \mathcal{F}, P)$.
Let $\tilde{P}$ be another equivalent probability measure on $(\Omega, \mathcal{F})$.
Let Z be almost surely positive random variable with $\mathrm{E}(Z) = 1$ defined by:

$$\tilde{P}(A) = \int_A Z(w) dP(w)$$

Then Z is called Radon-Nikodym derivative of $\tilde{P}$ with respect to P and it can be written:

$$Z(w) = \frac{d\tilde{P}(w)}{dP(w)}$$

Properties

1. $\mathrm{E}(Z) = 1$

   $\mathrm{E}(Z) = \int_\Omega Z(w) dP(w) = \tilde{P}(\Omega) = 1$

2. $\tilde{\mathrm{E}}(X) = \mathrm{E}(ZX)$ and $\tilde{\mathrm{E}}(\frac{1}{Z}X) = \mathrm{E}(X)$

   $\tilde{\mathrm{E}}(X) = \int_\Omega X(w) d\tilde{P}(w) = \int_\Omega X(w) \cdot \frac{dP(w)}{dP(w)} d\tilde{P}(w) = \int_\Omega X(w) Z(w) dP(w) = \mathrm{E}(ZX)$

To get the idea how Z can be derived and applied to random variables we will consider the following example.

Example (S.Shreve)

Let X be standard normal variable on $(\Omega, \mathcal{F}, P)$. (i.e. X $\sim N(0,1)$)
Let $Y = X + \theta$, where θ is some constant. We want to find Z and measure $\tilde{P}$ such that Y $\sim N(0,1)$) under $\tilde{P}$.

Derivation, non-strict

For a set A that contains $\bar{w}$ and small enough the following approximately holds:

$$Z(\bar{w}) \approx \frac{\tilde{P}(A)}{P(A)}$$

We use this observation and denote $x = X(\bar{w})$ and $y = Y(\bar{w}) = x + \theta$.

1. At first, let $B(x, \epsilon) = [x - \frac{\epsilon}{2}, x + \frac{\epsilon}{2}]$. Then by midpoint rule:

$$P(X \in B(x,\epsilon)) = \int_{x-\frac{\epsilon}{2}}^{x+\frac{\epsilon}{2}} \frac{1}{\sqrt{2\pi}} e^{\frac{-u^2}{2}} du = \epsilon \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{-X(\bar{w})^2}{2}}$$

2. Secondly, we know that under $\tilde{P}$-measure Y will have standard normal distribution. Therefore, we denote $B(y, \epsilon) = [y - \frac{\epsilon}{2}, y + \frac{\epsilon}{2}]$ and

$$\tilde{P}(Y \in B(y,\epsilon)) = \epsilon \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{-Y(\bar{w})^2}{2}}$$

3. Now we prove that $B(x, \epsilon)$ and $B(y, \epsilon)$ are the same sets and denote them $A(\bar{w}, \epsilon)$.

$$\{X \in [x - \frac{\epsilon}{2}, x + \frac{\epsilon}{2}]\} = \{X + \theta \in [x + \theta - \frac{\epsilon}{2}, x + \theta + \frac{\epsilon}{2}]\} =$$

$$= \{Y \in [y - \frac{\epsilon}{2}, y + \frac{\epsilon}{2}]\} = \{Y \in B(y, \epsilon)$$

$A(\bar{w}, \epsilon)$ is small when ϵ is small.

4. Finally,

$$Z(\bar{w}) \approx \frac{\tilde{P}(A)}{P(A)} = \frac{\epsilon \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{-Y(\bar{w})^2}{2}}}{\epsilon \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{-X(\bar{w})^2}{2}}} =$$

$$= e^{-\frac{Y(\bar{w}^2}{2} + \frac{X(\bar{w})^2}{2}} = e^{-x\theta - \frac{1}{2}\theta^2}$$

$$Z(w) = e^{-X(w)\theta - \frac{1}{2}\theta^2}$$

Properties

We check that $Z(w) > 0$ and $\mathrm{E}(Z) = 1$ and that Y indeed has standard normal distribution under $\tilde{P}$.

1. First is obvious because it's exponent.

2. $\mathrm{E}(Z) = \int_{-\infty}^{+\infty} e^{-x\theta - \frac{1}{2}\theta^2} \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} dx = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-(x+\theta)^2}{2}} d(x+\theta) = 1$

3. Let's check MGF for Y under $\tilde{P}$

   $\tilde{\mathrm{E}}(e^{tY}) = \mathrm{E}(Ze^{tY}) = \int_{-\infty}^{+\infty} e^{-x\theta - \frac{1}{2}\theta^2} \cdot e^{t(x+\theta)} \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} dx =$
   $= \int_{-\infty}^{+\infty} e^{t(x+\theta)} \frac{1}{\sqrt{2\pi}} e^{\frac{-(x+\theta)^2}{2}} d(x+\theta) = \int_{-\infty}^{+\infty} e^{t(y)} \frac{1}{\sqrt{2\pi}} e^{\frac{-(y)^2}{2}} dy =$
   $= e^{\frac{1}{2}t^2}$

   This means that under $\tilde{P}$, Y has standard normal distribution. In other words:

   $$\tilde{P}(Y \le b) = \int_{\{w:Y(w)\le b\}} d\tilde{P}(w) = \int_{\Omega} I_{\{w:Y(w)\le b\}} Z(w) dP(w) =$$

   $= \int_{\Omega} I_{\{w:X(w)+\theta\le b\}} Z(w) dP(w) = \int_{\Omega} I_{\{w:X(w)\le b-\theta\}} Z(w) dP(w) =$

   $= \int_{-\infty}^{b-\theta} e^{-x\theta - \frac{1}{2}\theta^2} \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} dx = \int_{-\infty}^{b-\theta} \frac{1}{\sqrt{2\pi}} e^{\frac{-(x+\theta)^2}{2}} dx = \int_{-\infty}^{b} \frac{1}{\sqrt{2\pi}} e^{\frac{-(y)^2}{2}} dy$

which can be rewritten as

$$\mathrm{Law}(Y|\tilde{P}) = N(0,1)$$

To make the above even more clear let's "meditate" on the result. The notation is not formal and only used for better understanding.

Let's consider $\bar{\omega}$ such that $X(\bar{\omega}) = 0$.
Then for standard probability measure P:

$$P\left(\bar{\omega} : X(\bar{\omega}) = 0\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{X(\bar{\omega})^2}{2}} = \frac{1}{\sqrt{2\pi}}$$

Since in the above $Y(w) = (X(w) + \theta)$
In standard probability measure P:

$$P(\bar{\omega} : Y(\bar{\omega}) = \theta) = P\left(\bar{\omega} : X(\bar{\omega}) = 0\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(Y(\bar{w})-\theta)^2}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{X(\bar{\omega})^2}{2}} = \frac{1}{\sqrt{2\pi}}$$

Since we changed the measure to $\tilde{P}$ such that:

$$\tilde{P}(Y = y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{(X(w)+\theta)^2}{2}}$$

For $\bar{\omega}$ such that $X(\bar{\omega}) = 0$ we now have:

$$\tilde{P}(Y(\bar{\omega}) = X(\bar{\omega}) + \theta = 0 + \theta = \theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(X(\bar{\omega})+\theta)^2}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{(\theta)^2}{2}}$$

We can see that in new measure $\tilde{P}$ less probability is assigned to $\bar{\omega}$ such that $X(\bar{\omega}) = 0$. The intuition is that since in measure P, $Y \sim N(\theta, 1)$ and we wanted $Y \sim N(0, 1)$ to be in measure $\tilde{P}$, we've "shifted" the distribution to the left.
I.e. previously peak of distribution was at $\bar{\omega}$ such that $X(\bar{\omega}) = 0$ and in new measure it is at $\tilde{\omega}$ such that $X(\tilde{\omega}) = -\theta$

Stock under Risk-Neutral Measure

One of the assumptions of Black-Scholes model is that stock follows GBM:

$$dS_t = \mu_t S_t dt + \sigma_t S_t dW_t$$

The problem is μ is not known for sure (and might depend on the attitude towards risk for different agents). The idea to change drift parameter

$$dS_t = \mu_t S_t dt + \sigma_t S_t dW_t = \mu_t S_t dt + \sigma_t S_t \left(dW_t + \frac{\mu_t - r_t}{\sigma_t} dt - \frac{\mu_t - r_t}{\sigma_t} dt \right) =$$

$$= \mu_t S_t dt - \sigma_t S_t \left(\frac{\mu_t - r_t}{\sigma_t} dt \right) + \sigma_t S_t (dW_t + \frac{\mu_t - r_t}{\sigma_t} dt) =$$

$$= r_t S_t dt + \sigma_t S_t (dW_t + \frac{\mu_t - r_t}{\sigma_t} dt)$$

The idea now is to introduce measure $\tilde{P}$ such that $d\tilde{W}_t = dW_t + \frac{\mu_t - r_t}{\sigma_t} dt$ is a Brownian motion.
(By integration, this is the same as $\tilde{W}(t) = W(t) + \int_0^t \frac{\mu_t - r_t}{\sigma_t} dt$)

Girsanov's theorem, one dimension

Let W_t, $0 \leq t \leq T$ be a Brownian motion on $(\Omega, \mathcal{F}, P)$ and $\mathcal{F}_t$ - filtration generated by this W_t. Let θ_t be an adapted process.
Define:

1. $Z(t) = e^{-\int_0^t \theta_u dW_u - \frac{1}{2} \int_0^t \theta_u^2 du}$
2. $\tilde{W}(t) = W(t) + \int_0^t \theta_u du$

Then $\mathrm{E}[Z(t)] = 1$ and $\tilde{W}_t$ is a Brownian motion under $\tilde{P}$

In order to prove the above we need to remind the definition of condition expectation and state Bayes Theorem for change of measure.

Conditional expectation (by S.Shreve)

Let $(\Omega, \mathcal{F}, P)$ be a probability space, let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$.
The conditional expectation of X given $\mathcal{G}$, E$[X|\mathcal{G}]$ is any random variable that satisfies:

1. (Measurability)

   E$[X|\mathcal{G}]$ is $\mathcal{G}$-measurable

2. (Partial averaging)

   $\int_A \mathrm{E}[X|\mathcal{G}](w)dP(w) = \int_A X(w)dP(w)$ for all A$\in \mathcal{G}$

It won't be proved here but E$[X|\mathcal{G}]$ always exists and it's unique.
Hint
$A = \{Y - Z\} > 0$ is in $\mathcal{G}$.

For future notation $Z = Z(T)$ and $Z(t) = E[Z|\mathcal{F}_t]$

Lemma 1
Let $0 \leq t \leq T$ and Y be an $\mathcal{F}_t$-measurable random variable. Then

$$\tilde{\mathrm{E}}[Y] = E[ZY] = E[E(ZY|\mathcal{F}_t)] = E[YE(Z|\mathcal{F}_t)] = E[YZ(t)]$$

Bayes theorem for change of measure

Let $0 \leq s \leq t \leq T$ and Y be an $\mathcal{F}_t$-measurable random variable. Then

$$\tilde{\mathrm{E}}[Y|\mathcal{F}(s)] = \frac{1}{Z(s)}E[YZ(t)|\mathcal{F}(s)]$$

To prove that the above is true we have to check that it satisfies the definition of conditional expectation

1. (Measurability)

   Clearly
   $\frac{1}{Z(s)}E[YZ(t)|\mathcal{F}(s)]$ is $\mathcal{F}(s)$-measurable

2. (Partial averaging)

   Check that
$$\int_A \frac{1}{Z(s)}E[YZ(t)|\mathcal{F}(s)d\tilde{P} = \int_A Yd\tilde{P}$$

Proof

$$\int_A \frac{1}{Z(s)}E[YZ(t)|\mathcal{F}(s)]d\tilde{P} = \tilde{E}[\mathcal{I}_A\frac{1}{Z(s)}E[YZ(t)|\mathcal{F}(s)]] = \tilde{E}[\mathcal{I}_AM] = E[\mathcal{I}_AMZ(s)]$$

$= E[\mathcal{I}_A\frac{1}{Z(s)}E[YZ(t)|\mathcal{F}(s)]Z(s)] = E[E[\mathcal{I}_A \cdot YZ(t)|\mathcal{F}(s)]] = E[\mathcal{I}_A \cdot YZ(t)] =$
$= \tilde{E}(\mathcal{I}_AY)) =$
$= \int_A Yd\tilde{P}$

Now we return back and prove Girsanov's theorem. The idea is to use Levy's theorem.

Proof
Girsanov's theorem

1. Clearly, $\tilde{W}(t)$ has continuous paths and $\tilde{W}(0) = W(0) + \int_0^0 \theta_u du = 0$

2. Furthermore, $[\tilde{W}(t), \tilde{W}(t)]_0^T = T$
since $\int_0^T \theta_u du$ contributes zero quadratic variation. This can also be shown by considering:
$$d\tilde{W}(t) \cdot d\tilde{W}(t) = (dW(t) + \theta(t)dt)^2 = dW(t)dW(t) = dt$$

3. Now we need to show that $\tilde{W}(t)$ under $\tilde{P}$ is a martingale.

   First, we show that Z(t) is a martingale under P:

   Let
   $X(t) = -\int_0^t \theta_u dW_u - \frac{1}{2}\int_0^t \theta_u^2 du$

   $f(X = x) = e^x$, $f_x(x) = e^x$, $f_{xx}(x) = e^x$

   $dX(t) = -\theta_t dW_t - \frac{1}{2}\theta_t^2 dt$
   $dX(t) \cdot dX(t) = \theta^2(t)dt$

   Then by Ito's formula:
$$dZ(t) = df(X(t)) = e^x dX(t) + \frac{1}{2} e^x dX(t)dX(t) = \; = e^{X(t)}\left(-\theta_t dW_t - \frac{1}{2}\theta_t^2 dt + \frac{1}{2}\theta_t^2\right)$$
   which results in
$$dZ(t) = -Z(t)\theta(t)dW(t)$$
   or by integrating
$$Z(t) = Z(0) - \int_0^t Z(u)\theta(u)dW(u)$$

This shows that $Z(t)$ is a martingale and $\mathrm{E}[Z(t)] = Z(0) = 1$.

Now we show that $\tilde{W}(t)Z(t)$ is a martingale under P. For this we once more consider Ito's product rule.

$$d(XY) = XdY + YdX + dXdY$$

or in our case:

$$d(\tilde{W}(t)Z(t)) = \tilde{W}(t)dZ(t) + Z(t)d\tilde{W}(t) + d\tilde{W}(t)dZ(t) =$$

$$= -\tilde{W}(t)Z(t)\theta(t)dW(t) + Z(t)dW(t) + Z(t)\theta(t)dt - dW(t)Z(t)\theta(t)dW(t) =$$

$$= \left(-\tilde{W}(t)Z(t)\theta(t) + Z(t)\right)dW(t) = \left(-\tilde{W}(t)\theta(t) + 1\right)Z(t)dW(t)$$

Because this has no dt term this is a martingale.
Now let $0 \leq s \leq t \leq T$ be given.
Using both lemmas and the above we get that $\tilde{W}(t)$ under $\tilde{P}$ is a martingale:

$$\tilde{E}[\tilde{W}(t)|\mathcal{F}(s)] = \frac{1}{Z(s)}E[\tilde{W}(t)Z(t)|\mathcal{F}(s)] = \frac{1}{Z(s)}\tilde{W}(s)Z(s) = \tilde{W}(s)$$

To sum up, we have that that $\tilde{W}(t)$ under $\tilde{P}$ is

1. a martingale that starts at 0
2. has continuous paths
3. $[\tilde{W}(t), \tilde{W}(t)]_0^T = T$

Then by Levy's theorem $\tilde{W}(t) = W(t) + \int_0^t \theta(s)ds$
is a Brownian motion under $\tilde{P}$

$$\mathrm{Law}(\tilde{W}|\tilde{P}) = N(0,t)$$

In context of GBM we can now write:

$$dS(t) = r(t)S(t)dt + \sigma(t)S(t)d\tilde{W}(t)$$

Application of Change of Measure

In order to apply Change of Measure principle for particular example, we briefly remind ourselves Reflection principle and its main result.

Reflection Equality
Let $W(t)$ be a standard Brownian Motion and

$$\tau_m = min\{t \geq 0 : W(t) = m\}$$

.
Then

$$P(\tau_m \leq t, W(t) \leq w) = P(W(t) \geq= 2m - w)$$

The idea is that for each path that hits level m before t and ends lower than level w, there is a reflected path that hits level m before t and ends higher than level $2m - w$. But since $2m - w \geq m$, this guarantees that such "reflected" trajectory has τ_m before t.

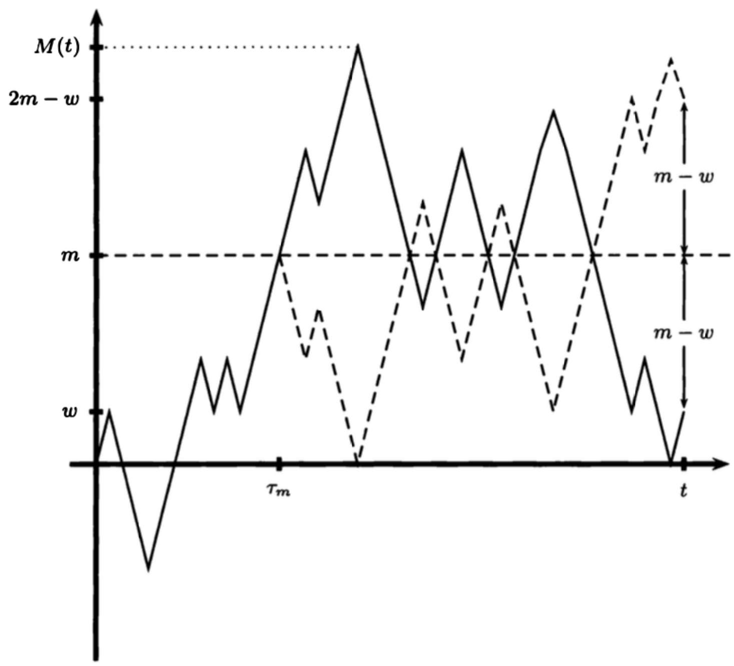

Figure 1: Brownian path and Reflected Brownian path

We further define

$$M(t) = \max_{0 \leq s \leq t} W(s)$$

We further observe that $\{\tau_m \leq t\}$ and $\{M(t) \geq m\}$ are equivalent as hitting barrier m before t guarantees that maximum $M(t)$ up to time t will be greater or equal then m.

This allows us to write

$$P(\tau_m \leq t, W(t) \leq w) = P(M(t) \geq m, W(t) \leq w) = P(W(t) \geq 2m - w)$$

This fact helps us to find joint distribution of $M(t)$ and $W(t)$. But first, we note that $M(t) \geq 0$ and $W(t) \leq M(t)$.
Therefore, range for pair of variables (M_t, W_t) is given by

$$\{(m, w) : w \leq m, m \geq 0\}$$

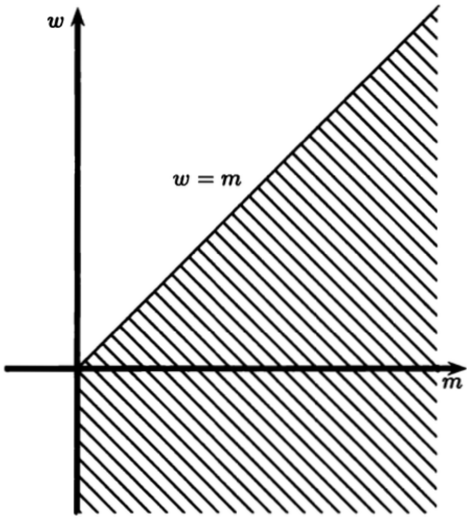

Figure 2: Range of M(t) and W(t)

Theorem
The joint density of $(M(t), W(t))$ is given by

$$f_{M(t),W(t)} = \frac{2(2m-w)}{t\sqrt{2\pi t}} e^{-\frac{(2m-w)^2}{2t}} \quad \text{for} \quad w \leq m, \quad m \geq 0$$

Proof
From one side

$$P(M(t) \geq m, W(t) \leq w) = \int_m^\infty \int_{-\infty}^w f_{M(t),W(t)}(x,y)dydx$$

On the other side

$$P(W(t) \geq 2m - w) = \int_{2m-w}^\infty \frac{1}{\sqrt{2\pi t}} e^{-\frac{z^2}{2t}} dz$$

By equaling two sides we get

$$\int_m^\infty \int_{-\infty}^w f_{M(t),W(t)}(x,y)dydx = \int_{2m-w}^\infty \frac{1}{\sqrt{2\pi t}} e^{-\frac{z^2}{2t}} dz$$

First, we take derivative with respect to m to get

$$-\int_{-\infty}^w f_{M(t),W(t)}(m,y)dy = -\frac{2}{\sqrt{2\pi t}} e^{-\frac{(2m-w)^2}{2t}}$$

Second, take derivative with respect to w to get

$$-f_{M(t),W(t)}(m,w) = -\frac{2(2m-w)}{t\sqrt{2\pi t}} e^{-\frac{(2m-w)^2}{2t}}$$

This leads to the desired result

$$f_{M(t),W(t)}(m,w) = \frac{2(2m-w)}{t\sqrt{2\pi t}} e^{-\frac{(2m-w)^2}{2t}}$$

Brownian motion with drift

The point of out interest now becomes process

$$\widehat{W}(t) = \widetilde{W}(t) + \alpha t$$

where on $(\Omega, \mathcal{F}, \widetilde{P})$

$$Law(\widetilde{W}(t)|\widetilde{P}) = N(0, t)$$

Note that under standard measure $\widetilde{P}$, $\widehat{W}(t)$ has a drift α.
We further define

$$\widehat{M}(t) = \max_{0 \leq s \leq t} \widehat{W}(s)$$

.
Again, we have range for pair of variables $(\widehat{M}_t, \widehat{W}_t)$ is given by

$$\{(m, w) : w \leq m, m \geq 0\}$$

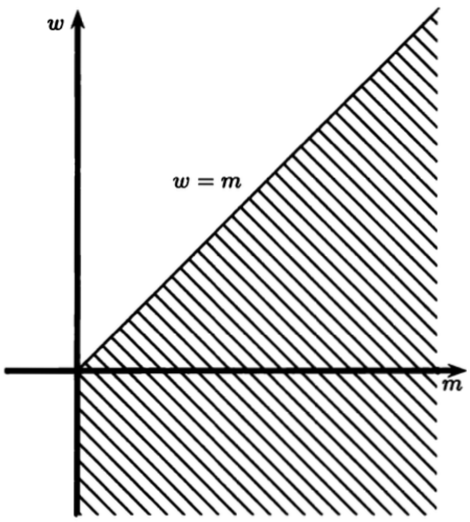

Figure 3: Range of $\widehat{M}(t)$ and $\widehat{W}(t)$

But how can one derive density for $\widehat{M}(t)$ under standard measure $\widetilde{P}$?

We recall Girsanov's theorem provided above and state the following Radon-Nikodym's derivative

$$\widehat{Z}(t) = e^{-\int_0^t \alpha d\widetilde{W}(s) - \frac{1}{2}\int_0^t \alpha^2 dt} = e^{-\alpha \widetilde{W}(t) - \frac{1}{2}\alpha^2 t} = e^{-\alpha \widehat{W}(t) + \frac{1}{2}\alpha^2 t}$$

This allows us to define new measure $\widehat{P}$ under which $\widehat{W}(t)$ has no drift

$$\widehat{P}(A) = \int_A Z(t) d\widetilde{P}$$

One can check that under $\widehat{P}$ (by MGF for example) that $\widehat{W}_t$ is standard Brownian motion with no drift.

$$Law(\widehat{W}_t | \widehat{P}) = N(0,t)$$

Therefore, under $\widehat{P}$-measure

$$\hat{f}_{\widehat{M}(t),\widehat{W}(t)} = \frac{2(2m-w)}{t\sqrt{2\pi t}} e^{-\frac{(2m-w)^2}{2t}} \quad \text{for} \quad w \leq m, \quad m \geq 0$$

To find out the density for $\{\widehat{M}(t), \widehat{W}(t)\}$ in $\widetilde{P}$ we refer to the Expectation property of Radon-Nikodym derivative.

$$\widetilde{P}(\widehat{M}(t) \leq m, \widehat{W}(t) \leq m) = \widetilde{E}[\mathbb{I}_{\widehat{M}(t) \leq m, \widehat{W}(t) \leq m)}] = \widehat{E}[\frac{1}{Z(t)} \mathbb{I}_{\widehat{M}(t) \leq m, \widehat{W}(t) \leq m)}] =$$

$$\widehat{E}[e^{\alpha \widehat{W}(t) - \frac{1}{2}\alpha^2 t} \mathbb{I}_{\widehat{M}(t) \leq m, \widehat{W}(t) \leq m)}] = \int_{-\infty}^{w} \int_{-\infty}^{m} e^{\alpha \widehat{W}(t) - \frac{1}{2}\alpha^2 t} f_{\widehat{M}(t),\widehat{W}(t)}(x,y) dy dx$$

Therefore, joint density under $\widetilde{P}$ is

$$\frac{\partial^2}{\partial w \partial m} \widetilde{P}(\widehat{M}(t) \leq m, \widehat{W}(t) \leq m) = e^{\alpha w - \frac{1}{2}\alpha^2 t} \cdot \hat{f}_{\widehat{M}(t),\widehat{W}(t)}(m,w) = e^{\alpha w - \frac{1}{2}\alpha^2 t} \cdot \frac{2(2m-w)}{t\sqrt{2\pi t}} e^{-\frac{(2m-w)^2}{2t}}$$

for

$$w \leq m, \quad m \geq 0$$

Otherwise, it is 0
One can derive $\widetilde{P}(\widehat{M}(t) \leq m)$ by taking integral over appropriate area.

Discrete Version of Girsanov's theorem

In order to reformulate continuous time theory for discrete time, we start by considering triple $(\Omega, \mathcal{F}, P)$ generated by 3-step Binomial model with the following parameters. For future notation $H = up$ and $T =$ down.

Model setup

1. $S_0 = 4$, $r = \frac{1}{4}$
2. $u = 2\ d = 1/2, \quad P(u) = 2/3 \quad P(d) = 1/3$
3. $S_n = S_0 u^k d^{(n-k)} = S_0 e^{X_1+X_2+X_3}$, where k - number of "ups", $n = 3$

   Note that this is the same as

4. $S_n = S_0 e^{X_1+X_2+X_3}$, where

$$X_i = \begin{cases} \ln(u) & p = \frac{2}{3} \\ \ln(d) & p = \frac{1}{3} \end{cases}$$

Then the following tree can be obtained

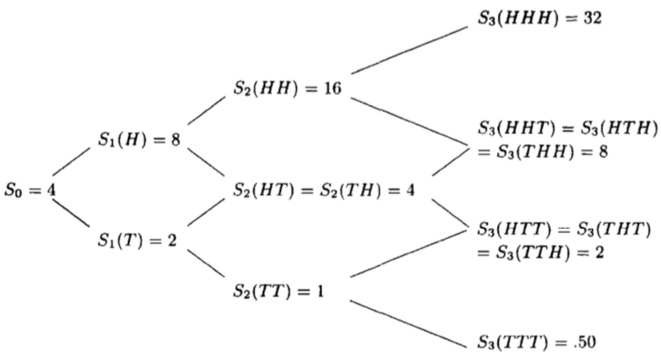

Figure 4: Three period model

with the following probabilities

$$\mathbb{P}(HHH) = \frac{8}{27},\quad \mathbb{P}(HHT) = \frac{4}{27},\quad \mathbb{P}(HTH) = \frac{4}{27},\quad \mathbb{P}(HTT) = \frac{2}{27},$$
$$\mathbb{P}(THH) = \frac{4}{27},\quad \mathbb{P}(THT) = \frac{2}{27},\quad \mathbb{P}(TTH) = \frac{2}{27},\quad \mathbb{P}(TTT) = \frac{1}{27}.$$

Figure 5: Probabilities under P

We now refer to Advanced statistics course and provide risk-neutral measure $\tilde{P}$ such that $S_n = \tilde{E}[\frac{S_(n+1)}{1+r}]$ i.e. it is a martingale.

$$\tilde{P}(u) = \frac{1+r-d}{u-d} = \frac{1}{2}$$

Then, the following probabilities are obtained for Binomial tree under $\tilde{P}$

$$\widetilde{\mathbb{P}}(HHH) = \frac{1}{8}, \quad \widetilde{\mathbb{P}}(HHT) = \frac{1}{8}, \quad \widetilde{\mathbb{P}}(HTH) = \frac{1}{8}, \quad \widetilde{\mathbb{P}}(HTT) = \frac{1}{8},$$
$$\widetilde{\mathbb{P}}(THH) = \frac{1}{8}, \quad \widetilde{\mathbb{P}}(THT) = \frac{1}{8}, \quad \widetilde{\mathbb{P}}(TTH) = \frac{1}{8}, \quad \widetilde{\mathbb{P}}(TTT) = \frac{1}{8}.$$

Figure 6: Risk-neutral Probabilities

$$\Omega = \{HHH, HHT, HTT...TTT\}$$

We now define Radon-Nykodym derivative for discrete time model and state it's properties

Radon-Nikodym Derivative

Let $P(w)$ and $\tilde{P}(w)$ be two probability measures on finite probability space such that $P(w) > 0$ and $\tilde{P}(w) > 0$ $\forall w \in \Omega$. Define $Z(w)$ by

$$Z(w) = \frac{\tilde{P}(w)}{P(w)}$$

and

$$\widetilde{P}(A) = \sum_{w \in A} Z(w)P(w) = \sum_{w \in A} \frac{\tilde{P}(w)}{P(w)}P(w) = \sum_{w \in A} \widetilde{P}(w)$$

.

Then $Z(w)$ is the Radon-Nikodym derivative with the following properties:

1. $P(Z > 0) = 1$
2. $E(Z) = 1$
3. $\widetilde{E}[X] = E[ZX]$

Proof

1. follows from the fact that $\tilde{P}(w) > 0 \quad \forall w \in \Omega$
2. follows from the following

$$E(Z) = \sum_{w \in \Omega} Z(w)P(w) = \sum_{w \in \Omega} \frac{\tilde{P}(w)}{P(w)} P(w) = \sum_{w \in \Omega} \tilde{P}(w) = 1$$

3. follows from the following

$$\widetilde{E}[X] = \sum_{w \in \Omega} X(w)\tilde{P}(w) = \sum_{w \in \Omega} X(w)\tilde{P}(w)\frac{P(w)}{P(w)} = \sum_{w \in \Omega} X(w)Z(w)P(w) = E[ZX]$$

Remark
For Binomial model the following definition of $Z(w)$ applies

$$Z(w) = \frac{\tilde{P}(w)}{P(w)} = \frac{\binom{n}{k}\tilde{p}^k \tilde{q}^{(n-k)}}{\binom{n}{k}p^k q^{(n-k)}} = \left(\frac{\tilde{p}}{p}\right)^k \left(\frac{\tilde{q}}{q}\right)^{n-k}$$

We now may define Radon-Nikodym derivative process

Radon-Nikodym derivative process

State $Z = Z_N$, where N is the number of steps in the Binomial Model.
Define

$$Z_n = E[Z|\mathcal{F}_n]$$

where $\mathcal{F}_n$ is filtration (information) up to moment n. Then Z_n is Radon-Nikodym derivative process with the property of martingale.

Proof
We need to show that $Z_n = E[Z_{n+1}|\mathcal{F}_n]$

$$E[Z_{n+1}|\mathcal{F}_n] = E[E(Z|\mathcal{F}_{n+1})|\mathcal{F}_n] = E(Z|\mathcal{F}_n) = Z_n$$

Regarding Figures above we can first compute $Z_3 = Z$ compute

$$Z(HHH) = \frac{27}{64}, \quad Z(HHT) = \frac{27}{32}, \quad Z(HTH) = \frac{27}{32}, \quad Z(HTT) = \frac{27}{16},$$
$$Z(THH) = \frac{27}{32}, \quad Z(THT) = \frac{27}{16}, \quad Z(TTH) = \frac{27}{16}, \quad Z(TTT) = \frac{27}{8}.$$

and then compute Radon-Nikodym derivative Process. For instance,

$$Z_2(HH) = E[Z|(HH)] = \frac{2}{3}Z(HHH) + \frac{1}{3}Z(HHT) = \frac{2}{3} \cdot \frac{27}{64} + \frac{1}{3} \cdot \frac{27}{32} = \frac{9}{16}$$

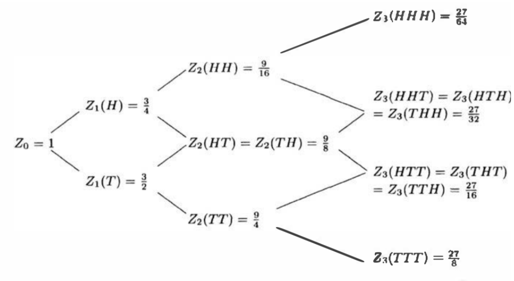

Figure 7: Radon-Nikodym Derivative Process

We now provide two supporting lemmas analogous to continuous time case.
Lemma 1
Let $0 \leq n \leq N$ and Y be an $\mathcal{F}_n$-measurable random variable. Then

$$\tilde{\mathrm{E}}[Y] = E[ZY] = E[E(ZY|\mathcal{F}_n)] = E[YE(Z|\mathcal{F}_n)] = E[YZ_n]$$

Lemma 2 (Bayes theorem for change of measure)

Let $0 \leq s \leq n \leq N$ and Y be an $\mathcal{F}_n$-measurable random variable. Then

$$\tilde{\mathrm{E}}[Y|\mathcal{F}_s] = \frac{1}{Z_s}E[YZ_n|\mathcal{F}_s]$$

Proof
One can prove lemma 2 by either using binomial distribution of $Z(w)$ given in remark above or considering formal definition as it was done in continuous time case.

We now proceed with generalized theory for discrete Girsanov's theorem and later return to the problem of Maximum distribution.

Girsanov's theorem: Discrete version

Consider the following iid process on $(\Omega, \mathcal{F}, P)$

$$h_n = \mu_n + \sigma_n \varepsilon_n$$

where μ_n and σ_n are predictable processes (i.e $\mathcal{F}_{n-1}$ measurable)

$$E[\mu_n | \mathcal{F}_{n-1}] = \mu_n$$

$$E[\sigma_n | \mathcal{F}_{n-1}] = \sigma_n$$

$$Law(\varepsilon_n; P) = N(0, 1)$$

This results in

$$Law(h_n | \mathcal{F}_{n-1}; P) = N(\mu_n, \sigma_n^2)$$

We further construct

$$H_n = h_1 + h_2 + ... + h_n = \sum_{k=1}^{n} h_k$$

Then

$$Law(H_n | \mathcal{F}_0; P) = N(\sum_{k=1}^{n} \mu_k, \sum_{k=1}^{n} \sigma_k^2)$$

We would like to find measure $\widetilde{P}$ such that it has no drift term

$$Law(H_n | \mathcal{F}_0; \widetilde{P}) = N(0, \sum_{k=1}^{n} \sigma_k^2)$$

The key role plays the following process

$$Z_n = e^{-\sum_{k=1}^{n} \frac{\mu_k}{\sigma_k}\varepsilon_k - \frac{1}{2}\sum_{k=1}^{n}\left(\frac{\mu_k}{\sigma_k}\right)^2} \qquad n \geq 1$$

Then Z_n is martingale and $E[Z_n] = 1$. We check this the following way

$$E[Z_n|\mathcal{F}_{n-1}] = E[Z_{n-1} \cdot e^{-\frac{\mu_n}{\sigma_n}\varepsilon_n - \frac{1}{2}\left(\frac{\mu_n}{\sigma_n}\right)^2}|\mathcal{F}_{n-1}] = Z_{n-1}E[\cdot e^{-\frac{\mu_n}{\sigma_n}\varepsilon_n - \frac{1}{2}\left(\frac{\mu_n}{\sigma_n}\right)^2}|\mathcal{F}_{n-1}] \overset{MGF}{=} Z_{n-1}$$

Furthermore

$$E[Z_1|\mathcal{F}_0] = 1$$

Therefore, we can use Z_n as Radon-Nikodym derivative to change measure. We state this by

$$Z_n(w) = \frac{d\widetilde{P}(w)}{dP(w)}$$

$$\widetilde{P}(A) = \int_A Z_n dP(w)$$

Now we are interested in distribution of h_n under $\widetilde{P}$. Consider $MGF(t)$ and apply Bayes theorem (Lemma 2)

$$\tilde{E}[e^{th_n}|\mathcal{F}_{n-1}] = \frac{1}{Z_{n-1}}E[e^{th_n}Z_n|\mathcal{F}_{n-1}] = \frac{1}{Z_{n-1}}E[e^{th_n}Z_{n-1}e^{-\frac{\mu_n}{\sigma_n}\varepsilon_n - \frac{1}{2}\left(\frac{\mu_n}{\sigma_n}\right)^2}|\mathcal{F}_{n-1}] =$$

$$= E[e^{t(\mu_n+\sigma_n\varepsilon_n)}e^{-\frac{\mu_n}{\sigma_n}\varepsilon_n - \frac{1}{2}\left(\frac{\mu_n}{\sigma_n}\right)^2}|\mathcal{F}_{n-1}] = e^{\mu_n t - \frac{1}{2}\left(\frac{\mu_n}{\sigma_n}\right)^2}E(e^{\varepsilon_n(t\sigma_n - \frac{\mu_n}{\sigma_n})}|\mathcal{F}_{n-1}) = e^{\frac{t^2\sigma_n^2}{2}}$$

We recall that this is MGF for standard normal variable with standard deviation σ_n and no drift. That is

$$Law(h_n|\mathcal{F}_{n-1};\widetilde{P}) = N(0,\sigma_n^2)$$

and

$$Law(H_n|\mathcal{F}_0;\widetilde{P}) = N(0,\sum_{k=1}^{n}\sigma_k^2)$$

Furthermore, if we assume that $\sigma = (\sigma_n)$ is a deterministic sequence we can obtain the following result

$$\tilde{E}[e^{tH_n}] = \tilde{E}[e^{t\sum_{k=1}^{N}h_k}] = \tilde{E}[e^{t\sum_{k=1}^{N-1}h_k} \cdot \tilde{E}[e^{th_n}|\mathcal{F}_{n-1}]] =$$

$$\tilde{E}[e^{t\sum_{k=1}^{N-1}h_k} \cdot \tilde{E}[e^{th_n}|\mathcal{F}_{n-1}]] = \tilde{E}[e^{\frac{t^2\sigma_n^2}{2}}e^{t\sum_{k=1}^{N-1}h_k}] = e^{\frac{t^2\sigma_n^2}{2}}\tilde{E}[e^{t\sum_{k=1}^{N-1}h_k}] =$$

$$= \text{by continuing the procedure} = e^{\sum_{k=1}^{N} \frac{t^2\sigma_n^2}{2}}$$

This confirms that under $\tilde{P}$, H_n has a normal distribution without drift term.

One way to non-strictly derive Radon-Nikodym derivative is by considering the following procedure for $h_n = \mu_n + \sigma_n \varepsilon_n$
Under $h_n \sim N(\mu_n, \sigma_n^2)$

$$dP(h_n = h) = \frac{1}{\sqrt{2\pi\sigma_n^2}} e^{-\frac{(h-\mu_n)^2}{2\sigma_n^2}}$$

Under $h_n \sim N(0, \sigma_n^2)$

$$d\widetilde{P}(h_n = h) = \frac{1}{\sqrt{2\pi\sigma_n^2}} e^{-\frac{(h)^2}{2\sigma_n^2}}$$

Then Z is $\frac{d\widetilde{P}}{dP}$ and we get

$$Z = \frac{\frac{1}{\sqrt{2\pi\sigma_n^2}} e^{-\frac{(h)^2}{2\sigma_n^2}}}{\frac{1}{\sqrt{2\pi\sigma_n^2}} e^{-\frac{(h-\mu_n)^2}{2\sigma_n^2}}}$$

Note that $h_n = h = \mu_n + \sigma_n \varepsilon_n$

$$Z = e^{-\frac{1}{2}\frac{\mu_n}{\sigma_n}\varepsilon_n - \frac{\mu_n^2}{\sigma_n^2}}$$

General result can be obtained by

$$Z_n = \frac{d\widetilde{P}}{dP}(h_1) \cdot \frac{d\widetilde{P}}{dP}(h_2) \ldots \cdot \frac{d\widetilde{P}}{dP}(h_n)$$

We've managed to provide Radon-Nikodym derivative for the process $H_n = \sum_1^n h_k$.

Another point of interest and martingale measure to find is for the process

$$S_n = S_0 e^{H_n}$$

The problem is that H_n being a martingale is not the same as e^{H_n} being a martingale. This arises due to the convexity of function and needs preliminary steps to be applied.

Consider

$$S_n = S_0 e^{h_1+h_2+...+h_n} = S_0 e^{h_1} e^{h_2} \cdot ... \cdot e^{h_n}$$

$$ln(\frac{S_n}{S_{n-1}}) = h_n$$

$$ln(1 + \frac{\Delta S_n}{S_{n-1}}) = h_n \quad \text{where} \quad \Delta S_n = S_n - S_{n-1}$$

Denote

$$\frac{\Delta S_i}{S_{i-1}} = \hat{h}_i$$

$$1 + \frac{\Delta S_i}{S_{i-1}} = 1 + \hat{h}_i = e^{h_i}$$

Let

$$\widehat{H}_n = \sum_1^n \hat{h}_i$$

Then

$$S_n = S_0 \prod_1^n (1 + \hat{h}_i) = S_0 \prod_1^n (1 + \Delta \hat{H}_i)$$

or

$$S_n = S_0 e^{\hat{H}_t} \prod_1^n (1 + \Delta H_i) e^{-\Delta \hat{H}_i} = S_0 \cdot \mathcal{E}(\hat{H})_n$$

$\mathcal{E}(\hat{H})_n$ is a stochastic sequence also called Doleans-Dade exponent which is a solution to

$$\Delta \mathcal{E}(\hat{H})_n = \mathcal{E}(\hat{H})_{n-1} \cdot \Delta \hat{H}_n$$

and used in pricing for Jump Processes

Proposition
For the S_n to be a martingale it is sufficient for the sequence $(\hat{H})_n$ to be a martingale. That is

$$E(\hat{H}_n|\mathcal{F}_{n-1}) = \hat{H}_{n-1}$$

or

$$E(\Delta\hat{H}_n|\mathcal{F}_{n-1}) = 0$$

This can be seen by the following composition

$$S_n = S_0 \prod_1^n (1 + \Delta\hat{H}_i) = S_{n-1} \cdot (1 + \Delta\hat{H}_n)$$

since then

$$E[S_n|\mathcal{F}_{n-1}] = S_{n-1} + S_{n-1}E[\Delta\hat{H}_n|\mathcal{F}_{n-1}]$$

We also note that

$$\Delta\hat{H}_n = \hat{h}_n = 1 + \hat{h}_n - 1 = e^{\Delta H_n} - 1$$

This gives us the following martingale condition

$$E[e^{\Delta H_n}|\mathcal{F}_{n-1}] = 1$$

or

$$E[e^{h_n}|\mathcal{F}_{n-1}] = 1$$

For the case $h_n = \mu_n + \sigma_n\varepsilon_n$

$$E[e^{\mu_n+\sigma_n\varepsilon_n}|\mathcal{F}_{n-1}] = 1$$

$$E[e^{\sigma_n\varepsilon_n}|\mathcal{F}_{n-1}] = e^{-\mu_n}$$

and by MGF gives us the following result

$$e^{\frac{1}{2}\sigma_n^2} == e^{-\mu_n}$$

or

$$\mu_n + \frac{1}{2}\sigma_n^2 = 0$$

(we want to choose such measure $d\tilde{P}$ and such that $\mu_n = -\frac{1}{2}\sigma_n^2$ in it)