MF921 Topics in Dynamic Asset Pricing

Stochastic Analysis & Stochastic Calculus in Quantitative Finance

Yuanhui Zhao

Boston University

Week 2

Part I, Chapter 15

Option Pricing via the Change of Numeraire Argument

Change of Numeraire: Motivation and Key Idea

In option pricing, we usually price under the risk-neutral measure using the money market account $B(t) = e^{rt}$ as the numeraire. But sometimes payoffs become simpler if we change the unit of measurement (the numeraire). Instead of measuring in "dollars," measure in "shares of stock".

The key idea is :

- Pick any strictly positive traded asset $N(t)$ as the numeraire.
- Then define a new probability measure $\tilde{\mathbb{P}}$ such that $\frac{S(t)}{N(t)}$ is a martingale under $\tilde{\mathbb{P}}$. No-arbitrage is preserved.

We first look at the details how this work (Radon Nikodym derivative & Girsanov Theorem) and then apply the scheme to price different type of options.

Change of Numeraire

Given $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}^*)$ with d-dim Brownian W:

- Money market account (baseline numeraire): $dB_t = B_t r_t dt$
- Traded asset $S(t)$: $dS_t = S_t(r_t dt + \sigma_t\, dW_t)$, $\quad \frac{S_t}{B_t}$ is a martingale.
- Derivative pricing rule: for payoff X_T at maturity T, $V_0 = \mathbb{E}^{\mathbb{P}^*}\left[\frac{X_T}{B_T}\right]$

Our goal is to pick another strictly positive traded asset $N(t)$ and define a new measure $\tilde{\mathbb{P}}$ such that $\frac{S(t)}{N(t)}$ is a martingale for every traded asset $S(t)$.

Change of Numeraire Con.

Oberve $\dfrac{S(t)}{B(t)}$ is a martingale under $\mathbb{P}^*$. We want $\dfrac{S(t)}{N(t)}$ to be a martingale under $\tilde{\mathbb{P}}$.
Define $\tilde{\mathbb{P}}$ via the Radon−Nikodym derivative with respect to $\mathbb{P}^*$:

$$\left.\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}^*}\right|_{\mathcal{F}_T} = Z_T := \frac{N(T)/B(T)}{N(0)/B(0)}$$

By construction, $\frac{N(T)}{B(T)}$ is a martingale under $\mathbb{P}^*$, $Z_T > 0$ and $\mathbb{E}^{\mathbb{P}^*}[Z_T] = 1$ and take any payoff X_T:

$$V(0) = N(0)\,\mathbb{E}^{\tilde{\mathbb{P}}}\left[\frac{X_T}{N(T)}\right] = N(0)\mathbb{E}^{\mathbb{P}^*}\left[\frac{X_T}{N(T)}\,Z_T\right] = \mathbb{E}^{\mathbb{P}^*}\left[\frac{X_T}{B(T)}\right]$$

So the choice of Radon−Nikodym derivative guarantees the prices are consistent under both measures and no arbitrage is preserved.

Change of Numeraire Con.

What is the $dS(t)$ looks like under meausre $\tilde{\mathbb{P}}$?

Note: Under Q, we have $\begin{cases} dS(t) = r(t)S(t)\,dt + \sigma(t)S(t)\,dW(t) \\ dN(t) = r(t)N(t)\,dt + \gamma(t)N(t)\,dW(t) \end{cases}$

Denote $\widehat{N}_t = \frac{N_t}{B_t}$, apply Itô we get $\frac{d\widehat{N}_t}{\widehat{N}_t} = \gamma_t\,dW_t$, $\widehat{N}_t = \widehat{N}_0 e^{\left(\int_0^t \gamma_s \cdot dW_s - \frac{1}{2}\int_0^t \|\gamma_s\|^2\,ds\right)}$.

Oberve that $Z_t = \frac{\widehat{N}_t}{\widehat{N}_0} = e^{\left(\int_0^t \gamma_s \cdot dW_s - \frac{1}{2}\int_0^t \|\gamma_s\|^2\,ds\right)}$

Girsanov's theorem says: if we define a new measure $\tilde{\mathbb{P}}$ via this Z_t, then the process

$$\tilde{W}(t) \;=\; W(t) - \int_0^t \gamma_s dt$$

is a Brownian motion under $\tilde{\mathbb{P}}$. Substitute into $dS(t)$ to get the $\tilde{\mathbb{P}}$ dynamics:

$$dS(t) = S(t)\left[(r(t) + \sigma(t)\cdot\gamma(t))\,dt + \sigma(t)\cdot d\tilde{W}(t)\right]$$

$$S(t) = S_0 \exp\left(\int_0^t \left(r(s) + \sigma(s)\cdot\gamma(s) - \tfrac{1}{2}\|\sigma(s)\|^2\right)ds + \int_0^t \sigma(s)\cdot d\tilde{W}(s)\right)$$

Black-Scholes Formula

Given r, σ are constant, we have $S(T) = S(0)\exp\left\{(r - \frac{1}{2}\sigma^2)T + \sigma W(T)\right\}$.

The no-arbitrage price for the call option:

$$\begin{aligned}
\psi_c(0) &= \mathbb{E}^{\mathbb{P}^*}(e^{-rT}(S(T) - K)^+) \\
&= \mathbb{E}^{\mathbb{P}^*}(e^{-rT}(S(T) - K)I(S(T) \geq K)) \\
&= \mathbb{E}^{\mathbb{P}^*}(e^{-rT}S(T)I(S(T) \geq K)) - Ke^{-rT}\mathbb{P}^*(S(T) \geq K) \\
&= I - Ke^{-rT} \cdot II
\end{aligned}$$

For II:

$$\begin{aligned}
II = \mathbb{P}^*(S(T) \geq K) = 1 - \Phi\left(\frac{\log(K/S(0)) - (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right) \\
= \Phi\left(\frac{\log(S(0)/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right)
\end{aligned}$$

Note: Φ is the CDF of the standard normal distribution.

Black-Scholes Formula Con.

For I, we apply the change of numeraire and use stock itself as numeraire. Then based on the eraly definition we have $\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}^*}\Big|_{\mathcal{F}_T} = Z_T := e^{-rT}\frac{S(T)}{S(0)}$ and $\gamma_t = \sigma$. Therefore, under $\tilde{\mathbb{P}}$ we have the following dynamics of $S(t)$:

$$\frac{dS_t}{S_t} = rdt + \sigma^2 dt + \sigma d\tilde{W}_t,\ S(t) = S(0)\exp\left\{(r+\sigma^2/2)t + \sigma\tilde{W}_t\right\}$$

Then we can rewrite I:

$$\begin{aligned} I = S(0)\mathbb{E}^{\mathbb{P}^*}\left(e^{-rT}\frac{S(T)}{S(0)}I(S(T)\geq K)\right) &= S(0)\mathbb{E}^{\tilde{\mathbb{P}}}(I(S(T)\geq K)) \\ &= S(0)\tilde{\mathbb{P}}(S(T)\geq K) \\ &= S(0)\Phi\left(\frac{\log(S(0)/K) + (r+\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right). \end{aligned}$$

Putting together, we have the price of the call option is given by:

$$I - Ke^{-rT}\cdot II = S(0)\Phi(d_+) - Ke^{-rT}\Phi(d_-)$$

where $d_\pm = \frac{\log(S(0)/K)+(r\pm\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$.

One Dimensional Barrier Options

Barrier options are path-dependent derivatives whose payoff is activated (knock-in) or extinguished (knock-out) if the underlying asset crosses a pre-specified barrier. They extend vanilla calls/puts by adding a barrier condition.
We first study continuously monitored barriers and derive Merton's closed-form pricing formulas (1973) for single-barrier options.
Given $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}^*)$ with 1-dim Brownian W. The Market setting following:

$$dB(t) = B(t)rdt \ , \ dS(t) = rS(t)dt + \sigma S(t)dW(t)$$

A continuously monitored barrier option has payoff $=$ vanilla option payoff $\times$ indicator of the barrier condition. For example:

- Up-and-out call:
$$V_0 = \mathbb{E}^{\mathbb{P}^*}\left[e^{-rT}(S(T)-K)^+ I\left\{\max_{0\leq t\leq T} S(t) \leq H\right\}\right], \quad H > S(0)$$
- Down-and-in put:
$$V_0 = \mathbb{E}^{\mathbb{P}^*}\left[e^{-rT}(K-S(T))^+ I\left\{\min_{0\leq t\leq T} S(t) \leq H\right\}\right], \quad H < S(0)$$

Study the case of the down-and-in call option (DAIC) with strike K, barrier $H < S(0)$:

$$\text{DAIC} = e^{-rT}\mathbb{E}^{\mathbb{P}^*}\left[(S(T)-K)^+ I\left\{\min_{0\leq t\leq T} S(t) \leq H\right\}\right]$$

One Dimensional Barrier Options Con.

For notation simplicity, denote a drifted Brownian motion:

$$W_{\mu,\sigma}(t) = \mu t + \sigma W(t), \quad M_t = \max_{0 \le s \le t} W_{\mu,\sigma}(s).$$

Some useful results from the reflection principle for a Brownian motion with a drift:
(i) When $x \le y$, $y > 0, \sigma > 0$:

- $P(W_{\mu,\sigma}(t) \le x, M_t \ge y) = e^{2\mu y/\sigma^2} \Phi\left(\frac{x-2y-\mu t}{\sigma\sqrt{t}}\right)$
- $P(W_{\mu,\sigma}(t) \le x, M_t \le y) = \Phi\left(\frac{x-\mu t}{\sigma\sqrt{t}}\right) - e^{2\mu y/\sigma^2} \Phi\left(\frac{x-2y-\mu t}{\sigma\sqrt{t}}\right)$

(ii) When $x \ge y > 0$, $\sigma > 0$:

- $P(W_{\mu,\sigma}(t) \le x, M_t \le y) = P(M_t \le y) = \Phi\left(\frac{y-\mu t}{\sigma\sqrt{t}}\right) - e^{2\mu y/\sigma^2} \Phi\left(\frac{-y-\mu t}{\sigma\sqrt{t}}\right)$
- $P(W_{\mu,\sigma}(t) \le x, M_t \ge y) = P(W_{\mu,\sigma}(t) \le x) - P(W_{\mu,\sigma}(t) \le x, M_t \le y) =$
$\Phi\left(\frac{x-\mu t}{\sigma\sqrt{t}}\right) - \Phi\left(\frac{y-\mu t}{\sigma\sqrt{t}}\right) + e^{2\mu y/\sigma^2} \Phi\left(\frac{-y-\mu t}{\sigma\sqrt{t}}\right)$

(iii) When $x \ge y$, $y < 0$, $\sigma > 0$:

- $P\left(W_{\mu,\sigma}(t) \ge x, \min_{0 \le s \le t} W_{\mu,\sigma}(s) \le y\right) = e^{2\mu y/\sigma^2} \Phi\left(\frac{-x+2y+\mu t}{\sigma\sqrt{t}}\right)$

One Dimensional Barrier Options Con.

Back to the valuation of DAIC:

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}^*}\left[e^{-rT}(S(T)-K)^+ I\left(\min_{0\le t\le T} S(t)\le H\right)\right] \\
&= \mathbb{E}^{\mathbb{P}^*}\left[e^{-rT}(S(T)-K) I\left(S(T)\ge K, \min_{0\le t\le T} S(t)\le H\right)\right] \\
&= \mathbb{E}^{\mathbb{P}^*}\left[e^{-rT}S(T) I\left(S(T)\ge K, \min_{0\le t\le T} S(t)\le H\right)\right] \\
&\quad - Ke^{-rT}P^*\left(S(T)\ge K, \min_{0\le t\le T} S(t)\le H\right) \\
&= I - Ke^{-rT}\cdot II
\end{aligned}
$$

For II:

$$
\begin{aligned}
II &= P^*\left(S(T)\ge K, \min_{0\le t\le T} S(t)\le H\right) \\
&= P\left\{W_{r-\frac{\sigma^2}{2},\sigma}(T)\ge \log(K/S(0)), \min_{0\le t\le T} W_{r-\frac{\sigma^2}{2},\sigma}(t)\le \log(H/S(0))\right\} \\
&= \exp\left\{\frac{2(r-\sigma^2/2)}{\sigma^2}\log(H/S(0))\right\}\cdot\Phi\left(\frac{2\log(H/S(0))-\log(K/S(0))+(r-\sigma^2/2)T}{\sigma\sqrt{T}}\right)
\end{aligned}
$$

For I, by changing of numeraire we can get:

$$
\begin{aligned}
I &= S(0)\mathbb{E}^{\mathbb{P}^*}\left(e^{-rT}\frac{S(T)}{S(0)}\cdot I\left\{S(T)\geq K, \min_{0\leq t\leq T} S(t)\leq H\right\}\right)\\
&= S(0)\tilde{P}\left(S(T)\geq K, \min_{0\leq t\leq T} S(t)\leq H\right)\\
&= S(0)P\left\{W_{r+\frac{\sigma^2}{2},\sigma}(T)\geq \log(K/S(0)), \min_{0\leq t\leq T} W_{r+\frac{\sigma^2}{2},\sigma}(t)\leq \log(H/S(0))\right\}\\
&= S(0)\cdot(H/S(0))^{\frac{2r}{\sigma^2}+1}\Phi\left(\frac{\log(\{H^2/S(0)\}/K)+(r+\sigma^2/2)T}{\sigma\sqrt{T}}\right)\\
&= (H/S(0))^{\frac{2r}{\sigma^2}-1}(H^2/S(0))\Phi\left(\frac{\log(\{H^2/S(0)\}/K)+(r+\sigma^2/2)T}{\sigma\sqrt{T}}\right)
\end{aligned}
$$

Putting the two terms together, we get $I - Ke^{-rT}\cdot II = (H/S(0))^{\frac{2r}{\sigma^2}-1}\mathsf{BSC}(H^2/S(0))$.
Where $\mathsf{BSC}(x)$ is the Black-Scholes formula for a call option with the initial stock price being x:

$$
\mathsf{BSC}(x) = x\Phi(d_+) - Ke^{-rT}\Phi(d_-) \text{ with } d_\pm = \frac{\log(x/K)+(r\pm\sigma^2/2)T}{\sigma\sqrt{T}}
$$

Exchange Options

Given $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}^*)$ with 2-dim independent Brownian, $W_1(t)$ and $W_2(t)$. We have two traded assets $S_1(t)$ and $S_2(t)$ with the following dynamics:

$$\frac{dS_1(t)}{S_1(t)} = rdt + \sigma_1 dW_1(t)$$
$$\frac{dS_2(t)}{S_2(t)} = rdt + \sigma_2 \left\{\rho dW_1(t) + \sqrt{1-\rho^2} dW_2(t)\right\}$$

The exchange option gives the holder the right, but not the obligation, to exchange asset S_2 for asset S_1 at maturity T. The price of this option as following:

$$\begin{aligned} u(0) &= \mathbb{E}^{\mathbb{P}^*}\left[e^{-rT}(S_1(T) - S_2(T))^+\right] \\ &= S_2(0)\mathbb{E}^{\mathbb{P}^*}\left[\frac{e^{-rT}S_2(T)}{S_2(0)}\left(\frac{S_1(T)}{S_2(T)} - 1\right)^+\right] \\ &= S_2(0)\mathbb{E}^{\tilde{\mathbb{P}}}\left[\left(\frac{S_1(T)}{S_2(T)} - 1\right)^+\right] \\ &= S_2(0)\mathbb{E}^{\tilde{\mathbb{P}}}\left[(F(T) - 1)^+\right] \end{aligned}$$

Exchange Options Con.

Apply Itô, we have the Radon−Nikodym derivative for numeraire:

$$\left.\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}^*}\right|_{\mathcal{F}_T} = Z_T^2 := \frac{e^{-rT}S_2(T)}{S_2(0)} = \exp\left[\sigma_2\left\{\rho W_1(T) + \sqrt{1-\rho^2}W_2(T)\right\} - \frac{T}{2}\sigma_2^2\right]$$

$$\left.\frac{d\hat{\mathbb{P}}}{d\mathbb{P}^*}\right|_{\mathcal{F}_T} = Z_T^1 := \frac{e^{-rT}S_1(T)}{S_1(0)} = \frac{e^{-rT}S_1(T)}{S_1(0)} = \exp\left(\sigma_1 W_1(T) - \frac{1}{2}\sigma_1^2 T\right)$$

By Girsanov theorem, under new measure $\tilde{\mathbb{P}}$:

$$\tilde{W}_1(t) = W_1(t) - \rho\sigma_2 t, \quad \tilde{W}_2(t) = W_2(t) - \sigma_2\sqrt{1-\rho^2}t$$

Apply Itô, we can get $d\ln S_1$, $d\ln S_2$:

$$\begin{aligned} d\ln F(t) &= d\ln S_1(t) - d\ln S_2(t) \\ &= \left[-\tfrac{1}{2}\sigma_1^2 - \tfrac{1}{2}\sigma_2^2 + \rho\sigma_1\sigma_2\right]dt + (\sigma_1 - \rho\sigma_2)d\tilde{W}_1 - \sigma_2\sqrt{1-\rho^2}d\tilde{W}\acute{c}_2. \end{aligned}$$

Apply Itô to $g(x) = e^x$ with $x = \ln F(t)$:

$$\frac{dF_t}{F_t} = d(\ln F_t) + \frac{1}{2}d < \ln F >_t = (\sigma_1 - \rho\sigma_2)d\tilde{W}_{1t} - \sigma_2\sqrt{1-\rho^2}d\tilde{W}_{2t}$$

Exchange Options Con.

Denote $\sigma = \sqrt{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2}$, $\tilde{W}(t) := \frac{1}{\sigma}\left\{(\sigma_1 - \rho\sigma_2)\tilde{W}_1(t) - \sigma_2\sqrt{1-\rho^2}\tilde{W}_2(t)\right\}$
Observe that $\tilde{W}$ is a standard Brownian motion under $\tilde{\mathbb{P}}$. We have $\frac{dF(t)}{F(t)} = \sigma d\tilde{W}(t)$, observe that $F_T = F_0 \exp\left(-\frac{1}{2}\sigma^2 T + \sigma\sqrt{T}Z\right), Z \sim N(0,1)$ under $\tilde{\mathbb{P}}$. Similarly, we have $F_T = F_0 \exp\left(\frac{1}{2}\sigma^2 T + \sigma\sqrt{T}Z\right), Z \sim N(0,1)$ under $\hat{\mathbb{P}}$.

Then we can rewrite $u(0)$:

$$
\begin{aligned}
u(0) &= S_2(0)\mathbb{E}^{\tilde{\mathbb{P}}}\left[(F(T)-1)^+\right] \\
&= S_2(0)\mathbb{E}^{\tilde{\mathbb{P}}}\left[(F(T)-1)I(F(T)>1)\right] \\
&= S_2(0)\left[\mathbb{E}^{\tilde{\mathbb{P}}}[F_T I\{F_T>1\}] - \tilde{\mathbb{P}}(F_T>1)\right] \\
&= S_2(0)\left[\mathbb{E}^{\mathbb{P}^*}[\frac{e^{-rT}S_2(T)}{S_2(0)}\frac{S_1(T)}{S_2(T)}I\{F_T>1\}] - \tilde{\mathbb{P}}(F_T>1)\right] \\
&= S_2(0)\left[\frac{1}{S_2(0)}\mathbb{E}^{\mathbb{P}^*}[\frac{e^{-rT}S_1(T)}{S_1(0)}S_1(0)I\{F_T>1\}] - \tilde{\mathbb{P}}(F_T>1)\right]
\end{aligned}
$$

Exchange Options Con.

$$
\begin{aligned}
&= S_2(0)\left[\frac{S_1(0)}{S_2(0)}\mathbb{E}^{\hat{\mathbb{P}}}[I\{F_T > 1\}] - \tilde{\mathbb{P}}(F_T > 1)\right] \\
&= S_1(0)\hat{\mathbb{P}}[I\{F_T > 1\}] - S_2(0)\tilde{\mathbb{P}}(F_T > 1) \\
&= S_1(0)\Phi(d_+) - S_2(0)\Phi(d_-)
\end{aligned}
$$

Where:

$$
d_\pm = \frac{\log(F(0)) \pm \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}} = \frac{\log(S_1(0)/S_2(0)) \pm \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}.
$$

- (i) If the second asset is cash, or $S_2(t) = Ke^{-r(T-t)}$, then the formula degenerates to the Black-Scholes formula.
- (ii) The hedging strategy is given by long $\Phi(d_+)$ shares of the first asset and short $\Phi(d_-)$ shares of the second asset.

Two-Dimensional Barrier Options

Suppose we have two Wiener processes, $X(t)$ and $Y(t)$, governed by the following dynamics

$$dX(t) = \mu_1 dt + \sigma_1 dW_1(t), \quad X(0) = 0, \quad \sigma_1 > 0,$$
$$dY(t) = \mu_2 dt + \sigma_2 \left\{ \rho dW_1(t) + \sqrt{1-\rho^2} dW_2(t) \right\}, \quad Y(0) = 0, \quad \sigma_2 > 0,$$

where the W_1 and W_2 are two independent standard Brownian motions.
For $b > 0$, consider the first passage time of the process $Y(t)$:

$$\tau_b^Y = \inf\{t \geq 0 : Y(t) = b > 0\}.$$

We shall prove that the joint distribution between $X(T)$ and the first passage time of $Y(t)$ is given by:

$$P(X(T) < a, \tau_b^Y > T) = P\left(X(T) < a, \max_{0 \leq t \leq T} Y(t) < b \right)$$

$$= \Phi_2\left(\frac{a - \mu_1 T}{\sigma_1 \sqrt{T}}, \frac{b - \mu_2 T}{\sigma_2 \sqrt{T}}; \rho \right) - e^{2\mu_2 b/\sigma_2^2} \Phi_2\left(\frac{a - \mu_1 T - 2\rho b \sigma_1/\sigma_2}{\sigma_1 \sqrt{T}}, \frac{-b - \mu_2 T}{\sigma_2 \sqrt{T}}; \rho \right)$$

where $b > 0$ and $\Phi_2(x, y; \rho)$ denotes the bivariate normal distribution given by

$$\Phi_2(x, y; \rho) = P(Z_1 \leq x, Z_2 \leq y),$$

with Z_1 and Z_2 being two standard normal random variables with correlation ρ.

Remark:

- Above equation holds for both $a \geq b$ and $a \leq b$, as long as $b > 0$. That's more general than the 1D reflection principle formulas which needed to be split into separate cases depending on $a \leq b$ or $a \geq b$.
- when $\rho = 1$, $\mu_1 = \mu_2 = \mu$, $\sigma_1 = \sigma_2 = \sigma$, the two dimensional case reduces to the one-dimensional case, as it becomes:

$$\begin{aligned}
&P\left(X(T) < a, \max_{0\leq t\leq T} X(t) < b\right) \\
&= \Phi_2\left(\frac{a-\mu T}{\sigma\sqrt{T}}, \frac{b-\mu T}{\sigma\sqrt{T}}; 1\right) - e^{2\mu b/\sigma^2}\Phi_2\left(\frac{a-\mu T-2b}{\sigma\sqrt{T}}, \frac{-b-\mu T}{\sigma\sqrt{T}}; 1\right) \\
&= P\left\{Z \leq \frac{a-\mu T}{\sigma\sqrt{T}}, Z \leq \frac{b-\mu T}{\sigma\sqrt{T}}\right\} - e^{2\mu b/\sigma^2} P\left\{Z \leq \frac{a-\mu T-2b}{\sigma\sqrt{T}}, Z \leq \frac{-b-\mu T}{\sigma\sqrt{T}}\right\} \\
&= P\left\{Z \leq \min\left\{\frac{a-\mu T}{\sigma\sqrt{T}}, \frac{b-\mu T}{\sigma\sqrt{T}}\right\}\right\} - e^{2\mu b/\sigma^2} P\left\{Z \leq \min\left\{\frac{a-\mu T-2b}{\sigma\sqrt{T}}, \frac{-b-\mu T}{\sigma\sqrt{T}}\right\}\right\}
\end{aligned}$$

Which incorporates two cases in one dimensional case.

Two-Dimensional Barrier Options Con.

Next we proof the formula of the joint distribution between $X(T)$ and the first passage time of $Y(t)$:
[Proof]
Consider the case of $\sigma_1 = \sigma_2 = 1$. Define a new process $V(t)$ to decouple X and Y:

$$V(t) := X(t) - \rho Y(t)$$

First check independence between V and Y:

$$\begin{aligned}
dV(t)dY(t) &= (dX(t) - \rho dY(t))dY(t) \\
&= \left((1-\rho^2)dW_1 - \rho\sqrt{1-\rho^2}dW_2\right) \cdot \left(\rho dW_1 + \sqrt{1-\rho^2}dW_2\right) \\
&= (1-\rho^2)\rho(dW_1)^2 - \rho(1-\rho^2)(dW_2)^2 \\
&= (1-\rho^2)\rho dt - (1-\rho^2)\rho dt = 0
\end{aligned}$$

Since $V(T) = X(T) - \rho Y(T)$, it is Gaussian. Its mean is:

$$\mathbb{E}[V(T)] = \mu_1 T - \rho\mu_2 T$$

Its variance is:

$$\begin{aligned}\mathsf{Var}(V(T)) &= \mathsf{Var}(X(T)) + \rho^2\mathsf{Var}(Y(T)) - 2\rho\mathsf{Cov}(X(T), Y(T)) \\ &= T + \rho^2 T - 2\rho^2 T = (1-\rho^2)T\end{aligned}$$

Thus:

$$V(T) \sim N((\mu_1 - \rho\mu_2)T, (1-\rho^2)T).$$

Incidentally, the same logic applying to two standard normal random variables with correlation ρ also leads to a representation for the bivariate normal distribution:

$$\Phi_2(\alpha, \beta; \rho) = \int_{z_2=-\infty}^{\beta} \int_{z_1=-\infty}^{\alpha} \frac{1}{\sqrt{1-\rho^2}} \varphi\left(\frac{z_1 - \rho z_2}{\sqrt{1-\rho^2}}\right) \varphi(z_2)\, dz_1\, dz_2.$$

Where $\varphi(\cdot)$ is the standard normal density function, $\varphi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$.

Now, in terms of $V(T)$, we can rewrite $P(X(T) < a, \tau_b^Y > T)$ as:

$$P(X(T) < a, \tau_b^Y > T)$$

$$= \int_{x=-\infty}^{a} \int_{y=-\infty}^{b} P(X(T) \in dx, Y(T) \in dy, \tau_b^Y > T)$$

Note:the transformation is linear with determinant 1 and the independence of V and Y

$$= \int_{x=-\infty}^{a} \int_{y=-\infty}^{b} P(V(T) \in dx - \rho dy) P(Y(T) \in dy, \tau_b^Y > T)$$

There are two terms inside the integrand. For the first term since $V(T)$ has a normal distribution with mean $\mu_1 T - \rho\mu_2 T$ and variance $(1-\rho^2)T$, we have:

$$P(V(T) \in dx - \rho dy) = \frac{1}{\sqrt{(1-\rho^2)T}} \varphi\left(\frac{x - \rho y - \mu_1 T + \rho\mu_2 T}{\sqrt{(1-\rho^2)T}}\right),$$

Two-Dimensional Barrier Options Con.

For the second term, recall the eraly result in one-dim, When $x \leq y$, $y > 0, \sigma > 0$:

$$P(W_{\mu,\sigma}(t) \leq x, M_t \leq y) = \Phi\left(\frac{x - \mu t}{\sigma\sqrt{t}}\right) - e^{2\mu y/\sigma^2}\Phi\left(\frac{x - 2y - \mu t}{\sigma\sqrt{t}}\right)$$

We have for all $y < b, b > 0$:

$$P(Y(T) \leq y, \tau_b^Y > T) = \Phi\left(\frac{y - \mu_2 T}{\sqrt{T}}\right) - e^{2\mu_2 b}\Phi\left(\frac{y - 2b - \mu_2 T}{\sqrt{T}}\right).$$

Differentiating the above equation yields:

$$P(Y(T) \in dy, \tau_b^Y > T) = \frac{1}{\sqrt{T}}\varphi\left(\frac{y - \mu_2 T}{\sqrt{T}}\right) - \frac{1}{\sqrt{T}}e^{2\mu_2 b}\varphi\left(\frac{y - 2b - \mu_2 T}{\sqrt{T}}\right).$$

Plugging the above two terms into:

$$\int_{x=-\infty}^{a}\int_{y=-\infty}^{b} P(V(T) \in dx - \rho dy)P(Y(T) \in dy, \tau_b^Y > T)$$

Two-Dimensional Barrier Options Con.

$$P(X(T) < a, \tau_b^Y > T) = I - II$$

where:

$$I = \int_{x=-\infty}^{a} \int_{y=-\infty}^{b} \frac{1}{\sqrt{(1-\rho^2)T}} \varphi\left(\frac{x - \rho y - \mu_1 T + \rho\mu_2 T}{\sqrt{(1-\rho^2)T}}\right) \frac{1}{\sqrt{T}} \varphi\left(\frac{y - \mu_2 T}{\sqrt{T}}\right) dydx,$$

$$II = e^{2\mu_2 b} \int_{x=-\infty}^{a} \int_{y=-\infty}^{b} \frac{1}{\sqrt{(1-\rho^2)T}} \varphi\left(\frac{x - \rho y - \mu_1 T + \rho\mu_2 T}{\sqrt{(1-\rho^2)T}}\right) \frac{1}{\sqrt{T}} \varphi\left(\frac{y - 2b - \mu_2 T}{\sqrt{T}}\right) dydx,$$

and $\varphi(x) = e^{-x^2/2}/\sqrt{2\pi}$ is the standard normal density function.

With $\tilde{x} = \frac{x-\mu_1 T}{\sqrt{T}}, \quad \tilde{y} = \frac{y-\mu_2 T}{\sqrt{T}}$, Then:

$$dx = \sqrt{T} d\tilde{x}, \quad dy = \sqrt{T} d\tilde{y}$$

$$x \le a \iff \tilde{x} \le \frac{a - \mu_1 T}{\sqrt{T}}, \quad y \le b \iff \tilde{y} \le \frac{b - \mu_2 T}{\sqrt{T}}$$

we have

$$I = \int_{x=-\infty}^{a} \int_{y=-\infty}^{b} \frac{1}{\sqrt{(1-\rho^2)T}} \varphi\left(\frac{\tilde{x}-\rho\tilde{y}}{\sqrt{(1-\rho^2)}}\right) \frac{1}{\sqrt{T}} \varphi(\tilde{y}) dy dx$$

$$= \int_{-\infty}^{\frac{a-\mu_1 T}{\sqrt{T}}} \int_{-\infty}^{\frac{b-\mu_2 T}{\sqrt{T}}} \frac{1}{\sqrt{(1-\rho^2)}} \varphi\left(\frac{\tilde{x}-\rho\tilde{y}}{\sqrt{(1-\rho^2)}}\right) \varphi(\tilde{y}) d\tilde{y} d\tilde{x}$$

By the conditional-Gaussian factorization, this integrand is exactly the joint pdf of a standard bivariate normal (Z_1, Z_2) with correlation ρ:

$$f_{Z_1,Z_2}(\tilde{x},\tilde{y}) = \frac{1}{\sqrt{1-\rho^2}} \varphi\left(\frac{\tilde{x}-\rho\tilde{y}}{\sqrt{1-\rho^2}}\right) \varphi(\tilde{y})$$

Hence the double integral is, by definition,

$$I = \Phi_2\left(\frac{a-\mu_1 T}{\sqrt{T}}, \frac{b-\mu_2 T}{\sqrt{T}}; \rho\right)$$

Similarly, with

$$\hat{x} = \frac{x - \mu_1 T - 2\rho b}{\sqrt{T}}, \quad \hat{y} = \frac{y - 2b - \mu_2 T}{\sqrt{T}}$$

simplifying the term II yields

$$\begin{aligned}
II &= \int_{x=-\infty}^{a} \int_{y=-\infty}^{b} \frac{1}{\sqrt{(1-\rho^2)T}} \varphi\left(\frac{\hat{x} - \rho\hat{y}}{\sqrt{(1-\rho^2)}}\right) \frac{1}{\sqrt{T}} e^{2\mu_2 b} \varphi(\hat{y}) dy dx \\
&= \int_{-\infty}^{\frac{a-\mu_1 T - 2\rho b}{\sqrt{T}}} \int_{-\infty}^{\frac{-b-\mu_2 T}{\sqrt{T}}} \frac{1}{\sqrt{(1-\rho^2)}} \varphi\left(\frac{\hat{x} - \rho\hat{y}}{\sqrt{(1-\rho^2)}}\right) e^{2\mu_2 b} \varphi(\hat{y}) dy dx \\
&= e^{2\mu_2 b} \Phi_2\left(\frac{a - \mu_1 T - 2\rho b}{\sqrt{T}}, \frac{-b - \mu_2 T}{\sqrt{T}}; \rho\right),
\end{aligned}$$

from which the result follows. The general case can be reduced to this particular case by letting:

$$\tilde{X}(t) = X(t)/\sigma_1, \quad \tilde{Y}(t) = Y(t)/\sigma_2,$$
$$\tilde{b} = b/\sigma_2, \quad \tilde{a} = a/\sigma_1, \quad \tilde{\mu}_1 = \mu_1/\sigma_1, \quad \tilde{\mu}_2 = \mu_2/\sigma_2.$$

Two-Dimensional Barrier Options Con.

Given the joint distribution between $X(T)$ and the first passage time of $Y(t)$ by :

$$P(X(T) < a, \tau_b^Y > T) = P\left(X(T) < a, \max_{0\le t\le T} Y(t) < b\right)$$

$$= \Phi_2\left(\frac{a-\mu_1 T}{\sigma_1\sqrt{T}}, \frac{b-\mu_2 T}{\sigma_2\sqrt{T}}; \rho\right) - e^{\frac{2\mu_2 b}{\sigma_2^2}} \Phi_2\left(\frac{a-\mu_1 T - 2\rho b\sigma_1/\sigma_2}{\sigma_1\sqrt{T}}, \frac{-b-\mu_2 T}{\sigma_2\sqrt{T}}; \rho\right)$$

Remark:
(i) Using the facts that $P(X(T) > a, \tau_b^Y > T) = P(-X(T) < -a, \tau_b^Y > T)$, that the correlation between $-X(t)$ and $Y(t)$ is $-\rho$,we can show that for $b > 0$(the following equation will use for next example to price of an up-and-out option):

$$P(X(T) > a, \tau_b^Y > T) = P\left(-X(T) < -a, \max_{0\le t\le T} Y(t) < b\right)$$

$$= \Phi_2\left(-\frac{a-\mu_1 T}{\sigma_1\sqrt{T}}, \frac{b-\mu_2 T}{\sigma_2\sqrt{T}}; -\rho\right) - e^{\frac{2\mu_2 b}{\sigma_2^2}} \Phi_2\left(-\frac{a-\mu_1 T - 2\rho b\sigma_1/\sigma_2}{\sigma_1\sqrt{T}}, \frac{-b-\mu_2 T}{\sigma_2\sqrt{T}}; -\rho\right)$$

(ii) Using the fact that $P(X(T) < a, \tau_{-b}^Y > T) = P(X(T) < a, \tau_b^{-Y} > T)$, that the correlation between $X(t)$ and $-Y(t)$ is $-\rho$, we can show that for $b > 0$:

$$P(X(T) < a, \tau_{-b}^Y > T) = P\left(X(T) < a, \min_{0\le t\le T} Y(t) > -b\right) = P\left(X(T) < a, \max_{0\le t\le T} (-Y(t)) < b\right)$$

$$= \Phi_2\left(\frac{a-\mu_1 T}{\sigma_1\sqrt{T}}, \frac{b+\mu_2 T}{\sigma_2\sqrt{T}}; -\rho\right) - e^{\frac{2\mu_2 b}{\sigma_2^2}} \Phi_2\left(\frac{a-\mu_1 T + 2\rho b\sigma_1/\sigma_2}{\sigma_1\sqrt{T}}, \frac{-b+\mu_2 T}{\sigma_2\sqrt{T}}; -\rho\right)$$

Two-Dimensional Barrier Options Con.

Let's calculate the price of an up-and-out call option, we have the following set up:

$$U_0 = e^{-rT}\mathbb{E}^{\mathbb{P}^*}\left[(S_1(T)-K)^+ I\left\{\max_{0\le t\le T} S_2(t)\le H\right\}\right], \quad S_i(t)=S_i(0)e^{X_i(t)}, X_1 = X, X_2 = Y$$

Under the risk-neutral measure $\mathbb{P}^*$,

$$dX(t) = \mu_1 dt + \sigma_1 dW_1(t), \quad \mu_1 = r - \frac{1}{2}\sigma_1^2,$$

$$dY(t) = \mu_2 dt + \sigma_2\left\{\rho dW_1(t) + \sqrt{1-\rho^2}dW_2(t)\right\}, \quad \mu_2 = r - \frac{1}{2}\sigma_2^2,$$

with W_1, W_2 independent.

Write the barrier level in log space

$$b := \log\frac{H}{S_2(0)}, \quad \text{and } a := \log\frac{K}{S_1(0)}$$

Then we have:

$$\{\max S_2(t)\le H\} = \left\{\max S_2(0)e^{Y(t)}\le H\right\} = \left\{\max_{0\le t\le T} Y(t)\le b\right\}$$

$$\{S_1(T)>K\} = \{S_1(0)e^{X(t)}>K\} = \{X(T)>a\}$$

$$
\begin{aligned}
U_0 &= e^{-rT}\mathbb{E}^{\mathbb{P}^*}\left[(S_1(T)-K)^+ I\left\{\max_{0\le t\le T} S_2(t)\le H\right\}\right] \\
&= e^{-rT}\mathbb{E}^{\mathbb{P}^*}\left[(S_1(T)-K) I\left\{S_1(T)>K \max_{0\le t\le T} S_2(t)\le H\right\}\right] \\
&= e^{-rT}\mathbb{E}^{\mathbb{P}^*}\left[S_1(T) I\left\{X(T)>a,\ \max_{0\le t\le T} Y(t)\le b\right\}\right] - Ke^{-rT}\mathbb{P}^*(X(T)>a,\ \max_{0\le t\le T} Y(t)\le b) \\
&= I - Ke^{-rT}\cdot II
\end{aligned}
$$

Apply Itô, we have the Radon−Nikodym derivative for numeraire:

$$
\left.\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}^*}\right|_{\mathcal{F}_T} = Z_T := \frac{e^{-rT}S_1(T)}{S_1(0)} = \exp\left[\sigma_1 W_1(T) - \frac{1}{2}\sigma_1^2 T\right]
$$

By Girsanov theorem, under new measure $\tilde{\mathbb{P}}$:

$$
\tilde{W}_1(t) = W_1(t) - \sigma_1 t,\ \ \tilde{W}_2(t) = W_2(t)
$$

And the drifts of X and Y under $\tilde{\mathbb{P}}$ become:

$$
\mu_1^{(1)} = \mu_1 + \sigma_1^2 = r + \frac{1}{2}\sigma_1^2, \quad \mu_2^{(1)} = \mu_2 + \rho\sigma_1\sigma_2 = r - \frac{1}{2}\sigma_2^2 + \rho\sigma_1\sigma_2
$$

Two-Dimensional Barrier Options Con.

For I, using change of numeraire and the formula in remark (i) :

$$
\begin{aligned}
I =& \mathbb{E}^{\mathbb{P}^*}\left[e^{-rT}S_1(T)I\left\{X(T) > a, \max_{0\le t\le T} Y(t) \le b\right\}\right] \\
=& \mathbb{E}^{\mathbb{P}^*}\left[\frac{e^{-rT}S_1(T)}{S_1(0)}S_1(0)I\left\{X(T) > a, \max_{0\le t\le T} Y(t) \le b\right\}\right] \\
=& S_1(0)\tilde{\mathbb{P}}\left\{X(T) > a, \max_{0\le t\le T} Y(t) \le b\right\}
\end{aligned}
$$

$$
= S_1(0)\left[\Phi_2\left(-\frac{a-\mu_1^{(1)}T}{\sigma_1\sqrt{T}}, \frac{b-\mu_2^{(2)}T}{\sigma_2\sqrt{T}}; -\rho\right) - e^{\frac{2\mu_2^{(2)}b}{\sigma_2^2}}\Phi_2\left(-\frac{a-\mu_1^{(1)}T-2\rho b\sigma_1/\sigma_2}{\sigma_1\sqrt{T}}, \frac{-b-\mu_2^{(2)}T}{\sigma_2\sqrt{T}}; -\rho\right)\right]
$$

For II, apply the formula in remakr(i) directly:

$$
\begin{aligned}
II =& \mathbb{P}^*(X(T) > a, \max_{0\le t\le T} Y(t) \le b) \\
=& \Phi_2\left(-\frac{a-\mu_1 T}{\sigma_1\sqrt{T}}, \frac{b-\mu_2 T}{\sigma_2\sqrt{T}}; -\rho\right) - e^{\frac{2\mu_2 b}{\sigma_2^2}}\Phi_2\left(-\frac{a-\mu_1 T-2\rho b\sigma_1/\sigma_2}{\sigma_1\sqrt{T}}, \frac{-b-\mu_2 T}{\sigma_2\sqrt{T}}; -\rho\right)
\end{aligned}
$$

Part II: Chapter 17

Introduction to Stochastic Calculus for Jump Processes

Counting Processes

A counting process $N(t)$: tracks the number of events up to time t. We have the following Key properties of any counting process:

- $N(t) \geq 0$
- Takes integer values
- Non-decreasing $(N(s) \leq N(t)$ if $s < t)$
- Increment $N(t) - N(s)$ counts events in $(s, t]$

To make analysis tractable, the following assumptions are typically imposed:

- Independent increments: the number of events occurring in disjoint time intervals is statistically independent
- Stationary increments: distribution of increments depends only on interval length, not location.

These two properties are also underpin the definition of Brownian motion.

Two Equivalent Definitions of Poisson Processes

First definition(I) of a Poisson process:
Counting process $\{N(t), t \geq 0\}$ with rate $\lambda > 0$ such that :

1. $N(0) = 0$
2. The process exhibits independent and stationary increments
3. For each $t \geq 0$, the random variable $N(t)$ follows a Poisson distribution:

$$\mathbb{P}[N(t) = n] = e^{-\lambda t}\frac{(\lambda t)^n}{n!} \quad \text{for } n = 0, 1, 2, \dots, \quad \mathbb{E}[N(t)] = \lambda t$$

Second definition(II) of a Poisson process:
Counting process $\{N(t), t \geq 0\}$ with rate $\lambda > 0$ such that :

1. $N(0) = 0$
2. The process exhibits independent and stationary increments
3. $\mathbb{P}[N(h) = 1] = \lambda h + o(h)$
4. $\mathbb{P}[N(h) \geq 2] = o(h)$

Two Equivalent Definitions of Poisson Processes Con.

We show the equivalence of the two definitions:
Proof:
(i) $I \Rightarrow II$
For small h, use Taylor expansion of the exponential:

$$\mathbb{P}(N(h) = 1) = e^{-\lambda h}(\lambda h) = (1 - \lambda h + \frac{1}{2}(\lambda h)^2 + o(h^2))(\lambda h) = \lambda h + o(h)$$

Similarly:

$$\mathbb{P}(N(h) \geq 2) = 1 - \mathbb{P}(0) - \mathbb{P}(1) = 1 - \left(1 - \lambda h + \frac{1}{2}(\lambda h)^2 + o(h^2)\right) - (\lambda h + o(h)) = o(h)$$

(ii) $II \Rightarrow I$(Some Intuition)
Partition $(0, t]$ into m small subintervals of length $h = t/m$ and define increments $X_i = N(ih) - N((i-1)h)$:

$$N(t) = \sum_{i=1}^{m} X_i$$

Let $m \to \infty$, by the small-interval conditions:

$$\mathbb{P}\{X_i = 1\} = \lambda h + o(h), \quad \mathbb{P}\{X_i = 0\} = 1 - \lambda h + o(h), \quad \mathbb{P}\{X_i \geq 2\} = o(h)$$

Two Equivalent Definitions of Poisson Processes Con.

So each X_i behaves like a Bernoulli(λh). due to the independent increments property, the X_i are mutually independent. It follows that $N(t)$ is approximately binomial with parameters m and $p = \lambda h$.

As $m \to \infty$ and $p \to 0$, the classical Poisson approximation to the binomial distribution implies that $N(t)$ converges in distribution to a Poisson random variable with rate

$$mp = m\lambda h = m\lambda \frac{t}{m} = \lambda t,$$

which precisely corresponds to the distribution given in equation of the requirement three in the first definition.

Remark: Powerful tool for modeling infrequent extreme events. In financial contexts, poisson processes can capture market shocks and discontinuities missed by continuous-path models. Important for pricing derivatives sensitive to jump risk.

Interarrival and Waiting Times

Interarrival time T_n is the time between the $(n-1)$st and nth event. Waiting time S_n is the time of the nth event:

$$S_n = \sum_{i=1}^{n} T_i$$

Key properties:

- $S_n \leq t \iff N(t) \geq n$. ($n$th event occurs by time $t \iff$ at least n arrivals by t).
- Alternative representation of counting process:

$$N(t) = \max\{n : S_n \leq t\} = \min\{n : S_{n+1} > t\}.$$

The Third Definition of Poisson Processes

Suppose the counting process $N(t)$ satisfies the second definition of a Poisson process. Demonstrate that the interarrival times $T_n, n \geq 1$, are independent exponential random variables with rate λ. Consequently, the expected value of the first interarrival time is $\mathbb{E}[T_1] = 1/\lambda$.

[Proof]
Npte that since $T_1 > t$ means that there is no event before time t. Therefore, we have

$$P(T_1 > t) = P(N(t) = 0) = e^{-\lambda t}.$$

Furthermore,

$$P(T_2 > t | T_1 = s) = P(\text{no events in } (s, s+t] | T_1 = s) = P(\text{no events in } (s, s+t]) = e^{-\lambda t},$$

Thus, conditioning on T_1 we have

$$P(T_2 > t) = \int_0^\infty P(T_2 > t | T_1 = s) f_{T_1}(s) ds = \int_0^\infty e^{-\lambda t} f_{T_1}(s) ds = e^{-\lambda t}$$

Therefore, T_2 has an exponential distribution with same rate λ and T_1 and T_2 are independent. Repeating the same argument, we can show T_3, T_4.....

The Third Definition of Poisson Processes Con.

Since the sum of independent and identically distributed exponential random variables follows a gamma distribution:

$$S_n = \sum_{i=1}^{n} T_i \sim \text{Gamma}(n, \lambda), \quad f_{S_n}(t) = \frac{\lambda e^{-\lambda t} (\lambda t)^{n-1}}{\Gamma(n)}$$

Third equivalent definition(III), poisson process can be defined via arrival times:

$$N(t) = \max\{n : S_n \leq t\} = \min\{n : S_{n+1} > t\} \qquad (*)$$

With $T_i \overset{iid}{\sim} \text{Exp}(\lambda)$.

Note: If $N(t)$ satisfies Definition II, then T_i are exponential $\Rightarrow (*)$. Conversely, if we build $N(t)$ from i.i.d. exponential interarrivals via $(*)$, then $N(t)$ has Poisson(λt) law $\Rightarrow$ Definition I.

Moreover, a Poisson process can be expressed as $N(t) = M(t) - 1$, where $M(t)$ corresponds to a special case of a first passage time process, defned as:

$$M(t) = \min\{n > 0 : S_n > t\}, \quad S_0 = 0, \quad t > 0$$

Compound Poisson Processes

Suppose $N(t)$ is a Poisson process with rate λ. Suppose $\{Y_i\}_{i\geq 1}$ are i.i.d. random variables with finite mean (μ_Y) and variance (σ_Y^2), independent of $N(t)$. Then the process:

$$X(t) = \sum_{i=1}^{N(t)} Y_i$$

is called a compound Poisson process.

Interpretation: at each jump of $N(t)$, we add a random amount Y_i. This models cumulative claims, shocks, or losses.

The conditional expectation is given by:

$$E\left[\sum_{i=1}^{N(t)} Y_i | N(t) = n\right] = E\left[\sum_{i=1}^{n} Y_i | N(t) = n\right] = E\left[\sum_{i=1}^{n} Y_i\right] = n\mu_Y$$

Consequently,

$$E\left[\sum_{i=1}^{N(t)} Y_i | N(t)\right] = N(t)\mu_Y,$$

and by the law of iterated expectations:

$$E\left[\sum_{i=1}^{N(t)} Y_i\right] = E\left\{E\left[\sum_{i=1}^{N(t)} Y_i | N(t)\right]\right\} = E[N(t)\mu_Y] = \mu_Y \lambda t.$$

Intuition: Think of $N(t)$ as "number of claims by time t" and Y_i as "size of each claim". Then $\sum_{i=1}^{N(t)} Y_i$ is total claim cost.The expected cost = (expected number of claims) $\times$ (expected claim size).

Compound Poisson Processes Con.

The conditional variance is given by:

$$\mathsf{Var}(\sum_{i=1}^{n} Y_i | N(t) = n) = \mathsf{Var}\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} \mathsf{Var}(Y_i) + \sum_{i=1}^{n} \sum_{j \neq i}^{n-1} \mathsf{Cov}(Y_i, Y_j) = n\sigma_Y^2$$

By law of total variance:

$$\begin{aligned}
\mathsf{Var}(X(t)) &= \mathbb{E}[\mathsf{Var}(X(t)|N(t))] + \mathsf{Var}(\mathbb{E}[X(t)|N(t)]) \\
&= \mathbb{E}[N(t)\sigma_Y^2] + \mathsf{Var}(N(t)\mu_Y) \\
&= \sigma_Y^2 \mathbb{E}[N(t)] + \mu_Y^2 \mathsf{Var}(N(t)) \\
&= \lambda t(\sigma_Y^2 + \mu_Y^2) = \lambda t \mathbb{E}[Y^2]
\end{aligned}$$

Compound Poisson Processes Con.

Determining the exact distribution function of a compound Poisson process is substantially more intricate. However, one can effectively compute its Laplace transform. Specifically, let $\psi(\theta)$ denote the Laplace transform of the claim size distribution Y, i.e., $\psi(\theta) = E[e^{-\theta Y}]$. The Laplace transform of the compound Poisson process $\sum_{i=1}^{N(t)} Y_i$ is given by:

$$
\begin{aligned}
E\left[\exp\left\{-\theta \sum_{i=1}^{N(t)} Y_i\right\}\right] &= E\left[E\left[\exp\left\{-\theta \sum_{i=1}^{N(t)} Y_i\right\}\right] \middle| N(t)\right] \\
&= E\left[\{\psi(\theta)\}^{N(t)}\right] \\
&= \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} (\psi(\theta))^n \\
&= e^{-\lambda t} \exp\{\lambda t \psi(\theta)\}.
\end{aligned}
$$

A Data Set

Consider the following data set from Biihlman (1970, p. 107, Mathematical Methods in Risk Theory, Springer), which is based on accident claims in 1961 for a class of automobile insurance by firms in Switzerland.

Number of Claims	Obs. Frequencies
0	103,704
1	14,075
2	1,766
3	255
4	45
5	6
6	2

To see whether the Poisson distribution fits the data, we shall fit the data first under Poisson distribution with the unknown parameter λ estimated by using Maximum Likelihood Estimation.

A Data Set Con.

Consider i.i.d. samples A_i, $i = 1, 2, ..., n$, from a Poisson distribution with rate λ. We want to estimate λ. The likelihood is given by:

$$L(\lambda) = \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{A_i}}{A_i!}.$$

Therefore, the log-likelihood is:

$$\log L(\lambda) = \sum_{i=1}^{n} \{-\lambda + A_i \log(\lambda) - \log(A_i!)\} = -n\lambda + \log(\lambda) \sum_{i=1}^{n} A_i - \sum_{i=1}^{n} \log(A_i!)$$

Taking derivatives with respect to λ and then setting them to zero yield:

$$-n + \frac{\sum_{i=1}^{n} A_i}{\lambda} = 0$$

Therefore, the estimators for λ is $\hat{\lambda} = \frac{1}{n} \sum\limits_{i=1}^{n} A_i$.

A Data Set Con.

In our example of the accident claims in 1961 for a class of automobile insurance by firms in Switzerland, we have:

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} A_i$$

$$= \frac{1}{n} \{0 \times 103,704 + 1 \times 14,075 + 2 \times 1,766 + 3 \times 255 + 4 \times 45 + 5 \times 6 + 6 \times 2\}$$

$$= 0.15514$$

where the total sample size is:

$$n = 103704 + 14075 + 1766 + 255 + 45 + 6 + 2 = 119853$$

Therefore, the fitted model for zero claims is:

$$n \times \left(e^{-\hat{\lambda}} \frac{\hat{\lambda}^0}{0!} \right) = 119853 \times e^{-\hat{\lambda}} = 119853 \times e^{-0.15514} = 102629.6$$

A Data Set Con.

We can continue to get all the fitted values for one claim, two claims, etc. The fitted model can be summarized in the following table.

Number of Claims	Obs. Frequencies	Fitted Poisson Model
0	103,704	102,629.6
1	14,075	15,922.0
2	1,766	1,235.1
3	255	63.9
4	45	2.5
5	6	0.1
6	2	0.0

Note that the Poisson process fts the tail part of the data rather poorly. Next, we shall discuss how to fit models to data better using an alternative model.

Pólya counting process

To get a better fit of the data, we shall consider the Pólya counting process, which is a generalization of the Poisson process.

The Pólya counting process $N(t)$ has a negative binomial distribution:

$$P(N(t) = i) = \binom{a+i-1}{i} \left(\frac{b}{t+b}\right)^a \left(\frac{t}{t+b}\right)^i, \quad t \geq 0, \quad i \geq 0, \quad a > 0, \quad b > 0$$

Note that $a > 0$ is not necessarily an integer, where the binomial coefficients for non-integers are defined as:

$$\binom{x}{y} = \frac{\Gamma(x+1)}{\Gamma(y+1)\Gamma(x-y+1)}$$

via the gamma function.

Pólya counting process Con.

The Pólya counting process is a generalization of Poisson processes, because it is a mixed Poisson process with the rate λ is a random variable Λ with a gamma density

$$\frac{b^a}{\Gamma(a)} e^{-bx} x^{a-1}$$

where $\Gamma(a)$ is a gamma function.

The properties of the Pólya counting process:

(1) Stationary but dependent increments. The Pólya counting process has a property that the arrival of one event tends to trigger more arrival events, leading to the positive correlation of increments. Indeed, $\text{Cov}(N(t),\ N(t+h) - N(t)) = ht\text{Var}(\Lambda) = ht\frac{a}{b^2}$.
[Proof]

$$\begin{aligned} E[N(t)(N(t+h) - N(t))] &= E[E[N(t)(N(t+h) - N(t))|\Lambda]] \\ &= E[t\Lambda \cdot h\Lambda] \\ &= th \cdot E[\Lambda^2] \end{aligned}$$

Pólya counting process Con.

Therefore:

$$
\begin{aligned}
&\mathsf{Cov}(N(t),\ N(t+h)-N(t)) \\
&= E[N(t)(N(t+h)-N(t))] - E[N(t)]E[(N(t+h)-N(t))] \\
&= thE[\Lambda^2] - thE[\Lambda]E[\Lambda] = th\mathsf{Var}(\Lambda)
\end{aligned}
$$

Due to dependent increments, it is more challenging to analyze the Pólya counting process.

(2) The Pólya counting process is a pure birth process with a birth rate being:

$$q_{i,i+1}(t) = \frac{a+i}{b+t},$$

where a and b are two parameters in the negative binomial distribution. This fact may be helpful when one tries to simulate the Pólya counting process.

Estimation

It is challenging to get the maximum likelihood estimators for a and b, involving infinite series and two implicit equations. However, it is easier to get estimators via the method of moments. Oberve that:

$$E\left[N(t)\right] = \frac{at}{b}, \quad \mathsf{Var}\left[N(t)\right] = a\frac{t}{b}\left(1 + \frac{t}{b}\right)$$

Setting up two equations

$$\frac{at}{b} = \bar{X}, \quad a\frac{t}{b}\left(1 + \frac{t}{b}\right) = S^2$$

yields

$$\hat{a} = \frac{(\bar{X})^2}{S^2 - \bar{X}}, \quad \hat{b} = \frac{t}{(S^2/\bar{X}) - 1}$$

Estimation Con.

In our example, $\sum X_i = 18594,\ \sum X_i^2 = 24376,\ n = 119853,\ t = 1$. So:

$$\bar{X} = 0.1551400466, \quad S^2 = 0.1793155390.$$

Thus:

$$\hat{a} = 0.9955716169, \quad \hat{b} = 6.417244540$$

Using the estimator $\hat{a}$ and $\hat{b}$ we get the following table for the negative binomial model.

Number of Claims	Obs. Frequencies	Fitted Poisson	Fitted Negative Binomial
0	103,704	102,629.6	103,760.8
1	14,075	15,922.0	13,927.1
2	1,766	1,235.1	1,873.5
3	255	63.9	252.2
4	45	2.5	34.0
5	6	0.1	4.6
6	2	0.0	0.6

The negative binomial model fits the data better than the Poisson model, especially in the tail part.

Introduction to Jump Diffusion Processes

A finite-activity jump process $J(t)$ is a process such that:

- (i) $J(t)$ is right-continuous. At a jump time t, $J(t-)$ is the value just before the jump, $J(t)$ is just after.
- (ii) There is no jump at time t, i.e. $J(t-) = J(t)$.
- (iii) Finite number of jumps in any finite interval. In other words, there is only a finite number of points such that $J(t-) \neq J(t)$.

These assumptions make stochastic calculus with jumps easier than with infinite-activity processes. Consider a stochastic process:

$$X(t) = X(0) + \int_0^t \theta(s)dW(s) + \int_0^t \mu(s)ds + J(t).$$

We can also write this as $X(t) = X^c(t) + J(t)$, where:

$$X^c(t) = X(0) + \int_0^t \theta(s)dW(s) + \int_0^t \mu(s)ds$$

is the part with a continuous sample path.

The Definition of Stochastic Integral

We can define the stochastic integral $\int_0^t \pi(s)dX(s)$. If $\pi(t) \in \mathcal{F}_t$ and $\pi(t)$ is left-continuous then the definition is given by

$$\int_0^t \pi(s)dX(s) = \int_0^t \pi(s)\theta(s)dW(s) + \int_0^t \pi(s)\mu(s)ds + \sum_{0<s\leq t} \pi(s)\Delta J(s),$$

or we can denote in differential form

$$\pi(t)dX(t) = \pi(t)\theta(t)dW(t) + \pi(t)\mu(t)dt + \pi(t)\Delta J(t).$$

Note that the sum is only active for the finite number of terms.

If $X(t)$ is a martingale, then under suitable integrability:

$$\mathbb{E}\left[\int_0^t \pi^2(s)\theta^2(s)\,ds\right] < \infty$$

The stochastic integral $\int_0^t \pi(s)\,dX(s)$ is also a martingale. $\pi(t)$ must be predictable (left-continuous). Otherwise, the martingale property may fail.

Itô Formula

Recall Itô formula without jumps, for a continuous semimartingale $X^c(t)$:

$$f(X^c(t)) = f(X^c(0)) + \int_0^t f'(X^c(s))\,dX^c(s) + \frac{1}{2}\int_0^t f''(X^c(s))\,d[X^c(s), X^c(s)]$$

The Itô formula for a jump process $X(t)$:

$$f(X(t)) = f(X(0)) + \int_0^t f'(X(s))dX^c(s) + \frac{1}{2}\int_0^t f''(X(s))d[X^c(s), X^c(s)] + \sum_{0<s\le t}[f(X(s)) - f(X(s-))]$$

Denote $\Delta f(X(s)) = f(X(s)) - f(X(s-))$.

The Itô formula in differential form:

$$df(X(t)) = f'(X(t))dX^c(t) + \frac{1}{2}f''(X(t))d[X^c(t), X^c(t)] + \Delta(f(X(t)))$$

Itô Formula Con.

Consider a jump-diffusion process for the return processes:

$$X(t) = X(0) + (\mu - \frac{1}{2}\sigma^2)t + \sigma W(t) + \sum_{i=1}^{N(t)} \ln(V_i).$$

Then the stock price $S(t) = e^{X(t)}$ is given by:

$$S(t) = e^{X(t)} = S(0) \exp\left\{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)\right\} \prod_{i=1}^{N(t)} V_i$$

Now apply Itô formula to $e^{X(t)}$, we have:

$$\begin{aligned} dS(t) &= e^{X(t)} d\left\{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)\right\} + \frac{1}{2}\sigma^2 e^{X(t)} dt + \Delta S(t) \\ &= (\mu - \frac{1}{2}\sigma^2) S(t) dt + S(t)\sigma dW(t) + \frac{1}{2}\sigma^2 S(t) dt + \Delta S(t) \\ &= \mu S(t) dt + S(t)\sigma dW(t) + \Delta S(t) \end{aligned}$$

Itô Formula Con.

Example 1: Consider a jump-diffusion process for the return processes:

$$X(t) = X(0) + (\mu - \frac{1}{2}\sigma^2)t + \sigma W(t) + \sum_{i=1}^{N(t)} \ln(V_i).$$

Then the stock price $S(t) = e^{X(t)}$ is given by:

$$S(t) = e^{X(t)} = S(0) \exp\left\{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)\right\} \prod_{i=1}^{N(t)} V_i$$

Now apply Itô formula to $e^{X(t)}$, we have:

$$\begin{aligned} dS(t) &= e^{X(t)} d\left\{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)\right\} + \frac{1}{2}\sigma^2 e^{X(t)} dt + \Delta S(t) \\ &= (\mu - \frac{1}{2}\sigma^2)S(t)dt + S(t)\sigma dW(t) + \frac{1}{2}\sigma^2 S(t)dt + \Delta S(t) \\ &= \mu S(t)dt + S(t)\sigma dW(t) + \Delta S(t) \end{aligned}$$

Itô Formula Con.

Now we need to compute $\Delta S(t)$. Suppose the ith jump occur at time t, the change in S is from $S(t-)$ to $S(t-)V_i$. Therefore, the change is $S(t-)V_i - S(t-) = S(t-)(V_i - 1)$. In this case, at the ith jump time t is $V_i - 1$, which can also be written more generally as

$$V_i - 1 = d\left(\sum_{i=1}^{N(t)}(V_i - 1)\right)$$

Which interpret as increment of the cumulative process at time t is equal to the jump size at t. Therefore, without referring to the specific jump time, we have:

$$\Delta S(t) = S(t-)d\left(\sum_{i=1}^{N(t)}(V_i - 1)\right)$$

Itô Formula Con.

Since in continuous parts of the trajectory (no jump), $S(t) = S(t-)$. In summary, we have:

$$
\begin{aligned}
dS(t) &= \mu S(t)dt + S(t)\sigma dW(t) + S(t-)d\left(\sum_{i=1}^{N(t)}(V_i - 1)\right) \\
&= \mu S(t-)dt + S(t-)\sigma dW(t) + S(t-)d\left(\sum_{i=1}^{N(t)}(V_i - 1)\right)
\end{aligned}
$$

or

$$
\frac{dS(t)}{S(t-)} = \mu dt + \sigma dW(t) + d\left(\sum_{i=1}^{N(t)}(V_i - 1)\right)
$$

Example 2: Suppose

$$\frac{dZ(t)}{Z(t-)} = \mu_0 dt + \sigma dW(t) + d\left[\sum_{i=1}^{N(t)} (V_i - 1)\right], \quad Z(0) = 1$$

where

$$\mu_0 = -\lambda E[V-1]$$

Then we shall show that $Z(t)$ is a martingale. We know from Example 1 that:

$$Z(t) = \exp\left\{-\frac{1}{2}\sigma^2 t + \sigma W(t)\right\} e^{\mu_0 t} \prod_{i=1}^{N(t)} V_i$$

But

$$E\left[\prod_{i=1}^{N(t)} V_i \middle| \mathcal{F}_s\right] = \prod_{i=1}^{N(s)} V_i \cdot E\left[\prod_{i=N(s)+1}^{N(t)} V_i\right]$$

Itô Formula Con.

However

$$
\begin{aligned}
E\left[\prod_{i=N(s)+1}^{N(t)} V_i\right] &= \sum_{n=0}^{\infty} \mathbb{E}\left[\prod_{i=1}^{n} V_i \middle| N(t)-N(s)=n\right]\cdot \mathsf{P}(N(t)-N(s)=n) \\
&= \sum_{n=0}^{\infty} e^{-\lambda(t-s)} \frac{(\lambda(t-s))^n}{n!} E\left[\prod_{i=0}^{n} V_i\right] \\
&= \sum_{n=0}^{\infty} e^{-\lambda(t-s)} \frac{(\lambda(t-s))^n}{n!} (E[V])^n \\
&= e^{-\lambda(t-s)} \sum_{n=0}^{\infty} \frac{(\lambda E[V](t-s))^n}{n!} \\
&= e^{-\lambda(t-s)} e^{\lambda E[V](t-s)} \\
&= e^{-\mu_0(t-s)}
\end{aligned}
$$

Itô Formula Con.

which show that for any $s < t$:

$$
\begin{aligned}
E[Z(t)|\mathcal{F}_s] &= E\left[\exp\left\{-\frac{1}{2}\sigma^2 t + \sigma W(t)\right\} e^{\mu_0 t} \prod_{i=1}^{N(t)} V_i \middle| \mathcal{F}_s\right] \\
&= Z(s) E\left[\exp\left\{-\frac{1}{2}\sigma^2 (t-s) + \sigma (W(t) - W(s))\right\} e^{\mu_0 (t-s)} \prod_{i=N(s)+1}^{N(t)} V_i \middle| \mathcal{F}_s\right] \\
&= Z(s) E\left[\exp\left\{-\frac{1}{2}\sigma^2 (t-s) + \sigma (W(t) - W(s))\right\}\right] \cdot E\left[e^{\mu_0 (t-s)} \prod_{i=N(s)+1}^{N(t)} V_i \middle| \mathcal{F}_s\right] \\
&= Z(s)
\end{aligned}
$$

Girsanov Theorem

If we use $Z(t)$ in Example 2, i.e.

$$\frac{dZ(t)}{Z(t-)} = \mu_0 dt + \sigma dW(t) + d\left[\sum_{i=1}^{N(t)} (V_i - 1)\right], \quad Z(0) = 1$$

to define a new probability measure P^* via $\frac{dP^*}{dP} = Z(t)$. Then Girsanov theorem says that under P^*:

- $W^*(t) = W(t) - \sigma t$ is a standard Brownian motion.
- The new jump density of V is given by $f_V^*(x) = \frac{1}{E[V]} x f_V(x)$.
- The new jump rate is given by $\lambda^* = \lambda E[V]$.

Instead of giving rigorous proof of the theorem, we present a heuristic derivation. First of all, $W^*(t)$ is a new Brownian motion due to the standard Girsanov theorem for Brownian motion. Thus, we shall focus on the jump part.

Girsanov Theorem Con.

Consider the jump arrival times $\tau_1, ..., \tau_n, ...,$ and the related jump sizes $V_1, V_2, ...,$. Conditioning on the event that $\{\tau_n \in [t, t+dt]\}$, we have the jump intensity under the old measure P is given by:

$$\lambda(dt, dx) := P(\tau_n \in [t, t+dt], V_n \in [x, x+dx] | \mathcal{F}_{t-}) = \lambda dt \cdot f_V(x) dx$$

This means that a jump happens in $[t, t+dt]$ with probability about λdt and conditional on a jump, the size distribution is $f_V(x)$. By the Bayes formula for the measure transform, we have the new intensity:

$$\begin{aligned}
\lambda^*(dt, dx) &= P^*(\tau_n \in [t, t+dt], V_n \in [x, x+dx] | \mathcal{F}_{t-}) \\
&= E^*[I(\tau_n \in [t, t+dt]) I(V_n \in [x, x+dx]) | \mathcal{F}_{t-}] \\
&= E\left[\frac{Z(t)}{Z(t-)} \cdot I(\tau_n \in [t, t+dt]) I(V_n \in [x, x+dx]) | \mathcal{F}_{t-}\right]
\end{aligned}$$

Conditioning on the event that $\{\tau_n \in [t, t+dt]\}$, $Z(t)$ is updated by a multiplicative factor V_n:

$$\frac{Z(t)}{Z(t-)} = V_n$$

Girsanov Theorem Con.

which yields:

$$\lambda^*(dt,dx) = E\left[E\left\{\frac{Z(t)}{Z(t-)}\cdot I(\tau_n\in[t,t+dt])I(V_n\in[x,x+dx])|\tau_n\in[t,t+dt]\right\}\middle|\mathcal{F}_{t-}\right]$$
$$= E\left[E\left\{V_nI(\tau_n\in[t,t+dt])I(V_n\in[x,x+dx])|\tau_n\in[t,t+dt]\right\}|\mathcal{F}_{t-}\right]$$
$$= E\left[I(\tau_n\in[t,t+dt])E\{V_nI(V_n\in[x,x+dx])|\tau_n\in[t,t+dt]\}|\mathcal{F}_{t-}\right]$$
$$= E\left[I(\tau_n\in[t,t+dt]|\mathcal{F}_{t-})\right]\cdot E\{V_nI(V_n\in[x,x+dx])|\mathcal{F}_{t-}\}$$
$$= \lambda dt\cdot E\left[V_nI(V_n\in[x,x+dx])|\mathcal{F}_{t-}\right],$$

via independence. Thus:

$$\lambda^*(dt,dx) = \lambda dt\cdot xf_V(x)dx = (E[V]\lambda)dt\cdot\frac{xf_V(x)}{E[V]}dx$$

which is the jump intensity of a new jump diffusion process with jump rate $\lambda^* = E[V]\lambda$, and jump size density:

$$\frac{xf_V(x)}{E[V]},\quad \text{with} \int_0^\infty \frac{xf_V(x)}{E[V]}dx = 1$$

where the normalizing constant $E[V]$ is needed to get a proper density function.