The Handbook of Mathematics, Statistics, and Various Other Topics

Thierry Bazier Matte

2015-2016

Contents

Part I

Machine Learning

- In **logistic regression**, the following model is often adopted:

$$
\begin{aligned}
P(y=1|x) &= \frac{1}{1+\exp(-\theta^T x)} = \sigma(\theta^T x), \\
P(y=0|x) &= 1-\sigma(\theta^T x)
\end{aligned}
$$

  with $\sigma(x)$ the **sigmoid function**.

- The **loss function** of the logistic regression is given by

$$
J(\theta) = -\sum_i (y^{(i)} \log(h(x^{(i)})) + (1-y^{(i)}) \log(1-h(x^{(i)}))).
$$

Part II

Applied Mathematics, Optimization and Risk Theory

1 Convexity and Risk Theory

- The **marginal utility** of utility u is defined by u''. Therefore, a decreasing marginal utility implies a risk-averse utility. [Eeckhoudt, p.9]

- **Jensen's Inequality.** If g is a convex function and X is a random variable, then

$$g(E[X]) \leq E[g(X)].$$

  Conversely, if g is concave, then

$$g(E[X]) \geq E[g(X)].$$

- The **Taylor expansion** of $E[f(X)]$ is given by:

$$E[f(X)] = f(E[X]) + \frac{1}{2}\operatorname{Var}[X] f''(E[X]) + \cdots + \frac{1}{j!}\gamma_j f^{(j)}(E[X]) + \cdots$$

  , where

$$\gamma_n = E[(X - E[X])^n].$$

- If g is convex and $g(0) \leq 0$, then g is **superadditive**, ie.

$$g(a) + g(b) \leq g(a+b).$$

  Conversely, if g is concave and $g(0) \geq 0$ then g is **subadditive**, ie.

$$g(a) + g(b) \geq g(a+b).$$

- The **certainty equivalent** of a lottery $\tilde{x}$ with utility u is defined by

$$CE(\tilde{x}) = u^{-1}(Eu(\tilde{x})).$$

Part III
Probability and Statistics

2 Basic Probability Rules

Let A and B be two events, possibly overlapping. Then,

- **Disjunction (or).**
$$P(A \vee B) = P(A) + P(B) - P(A \wedge B)$$
- **Conjunction (and).**
$$P(A \wedge B) = P(A) \times P(B|A)$$

3 Definitions

- The **cumulative distribution function** or **CDF** of a random variable is defined by
$$F_X(x) = P(X \leq x).$$
We also have the following identity:
$$P(a < X \leq b) = F_X(b) - F_X(a).$$
- If X is a continuous random variable, then its **probability density function** or **PDF** is defined by
$$f_X(x) = \frac{d}{dx} F_X(x).$$
Equivalently, we also have
$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt.$$
- By definition of the PDF,
$$P(a \leq X \leq B) = \int_a^b f_X(x)\, dx.$$
- The **quantile function** $Q : [0,1] \to R$, R being the range of X, is given, in the continuous and nice case by:
$$Q(p) = F_X^{-1}(p),$$
ie the value x such that
$$F_X(x) = P(X \leq x) = p,$$
and in the general case by
$$Q(p) = \inf\{x : p \leq F_X(x)\}.$$
- The quantile of q_p of $X \sim \mathcal{N}(\mu, \sigma^2)$ is obtained by using $(X - \mu)/\sigma$, ie. $\mu + \sigma q_p(\mathcal{N}(0,1))$.

4 Distribution Theory

- A **probability space** $(\Omega, \mathscr{F}, P)$ is a set of three parts:
  1. A sample space Ω of all possible outcomes;
  2. A set of events $\mathscr{F}$;
  3. An assignment of probability P for each of these events.
- The **expected value** of a random variable X with density f is defined by

$$E[X] = \int x\, f(x)\, dx.$$

In general, if X is defined on probability space $(\Omega, \mathscr{F}, P)$, then

$$E[X] = \int_\Omega X\, dP = \int_\Omega X(\omega) P(d\omega).$$

- The **law of total expectation** states that

$$E[X] = E[E[X|\mathscr{F}_t]].$$

- The **variance** of a random variable X is defined by

$$\begin{aligned} \mathrm{Var}(X) &= E[(X - E[X])^2] \\ \mathrm{Var}(X) &= E[X^2] - E[X]^2. \end{aligned}$$

- The **covariance** of two random variables is defined by

$$\begin{aligned} \mathrm{Cov}(X, Y) &= E[(X - E[X])(Y - E[Y])] \\ &= E[XY] - E[X]E[Y]. \end{aligned}$$

- The **correlation** of two random variables is defined by

$$\mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}(X)\,\mathrm{Var}(Y)}.$$

- The **covariance matrix** Σ of random variables X_i is defined by

$$\begin{aligned} \Sigma_{ij} &= \mathrm{Cov}(X_i, X_j) \\ \Sigma &= E[(X - E[X])(X - E[X])^T] \\ \Sigma &= \mathrm{diag}(\sigma) R\, \mathrm{diag}(\sigma). \end{aligned}$$

- The **correlation matrix** R of random variables X_i is defined by

$$\begin{aligned} R_i &= \mathrm{Corr}(X_i, X_j) \\ R &= \mathrm{diag}(\Sigma)^{-1/2} \Sigma\, \mathrm{diag}(\Sigma)^{-1/2}. \end{aligned}$$

- The **skewness** γ_1 of a random variable X is defined by

$$\gamma_1 = \frac{1}{\sigma^3} E[(X-\mu)^3].$$

- The **kurtosis** γ_2 of a random variable X is defined by

$$\gamma_2 = \frac{1}{\sigma^4} E[(X-\mu)^4].$$

- The **characteristic function** $\varphi_X(t)$ of a random variable X is defined by

$$\varphi_X(t) = E[\exp(itX)].$$

- The **moment generation function** $M_X(t)$ of a random variable X is defined by

$$M_X(t) = E[\exp(tX)].$$

  Let m_i be the i^{th} moment of X. Then,

$$M_X(t) = 1 + tm_1 + \tfrac{1}{2!}t^2 m_2 + \tfrac{1}{3!}t^3 m_3 + \cdots$$

- The **cumulants** κ_n of a random variable X are defined by

$$\kappa_n(X) = \frac{d^n}{dt^n} \log(M_X(t))\bigg|_{t=0}.$$

  The following rule holds

$$\kappa_n(aX + bY) = a^n \kappa_n(X) + b^n \kappa(Y).$$

  We also have

$$\begin{aligned}
\kappa_1(X) &= E[X] = \mu \\
\kappa_2(X) &= \mathrm{Var}(X) = \sigma^2 \\
\kappa_3(X) &= E[(X-\mu)^3] \\
\kappa_4(X) &= E[(X-\mu)^4] - 3\sigma^4.
\end{aligned}$$

  In particular, the skewness $\gamma_1 = \kappa_3/\sigma^3$ and the kurtosis $\gamma_2 = 3 + \kappa_4/\sigma^4$.

5 Stochastic Calculus

- A **brownian motion** has independant increments and $W_t \sim \mathcal{N}(0,t)$ (variance of t, so standard deviation $\sqrt{t}$).

- A d-dimensional brownian processes W is correlated with matrix R if

$$W_t - W_s \sim \mathcal{N}_d(0, R(t-s)).$$

- The **reflection principle** states that

$$P(M_t \geq a) = 2P(W_t \geq a),$$

  where

$$M_t = \sup_{0 \leq s \leq t} W_t.$$

6 Distributions

6.1 Normal Distribution

Notation	$\mathcal{N}(\mu, \sigma^2)$
Mean	μ
Variance	σ^2
Skewness	0
Kurtosis	3
PDF	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$
CDF	$\frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right)$
Quantile	$\mu + \sigma\sqrt{2}\operatorname{erf}^{-1}(2q-1)$
Quantile $\{0.01, 0.99\}$	± 2.32635
Quantile $\{0.025, 0.975\}$	± 1.95996

- Linear combination of normal variables remains normal:

$$\sum_i a_i X_{\mathcal{N}(\mu_i, \sigma_i^2)} \sim \mathcal{N}(\sum_i a_i \mu_i, \sum_i (a_i \sigma_i)^2).$$

6.2 Lognormal Distribution

Notation	$\mathcal{LN}(\mu, \sigma^2)$
Mean	$\exp(\mu + \sigma^2/2)$
Variance	$(\exp(\sigma^2) - 1)\exp(2\mu + \sigma^2)$

- By definition,

$$\exp X_{\mathcal{N}(\mu,\sigma^2)} \sim \mathcal{LN}(\mu, \sigma^2).$$

- The following properties hold:

$$aX_{\mathcal{LN}(\mu,\sigma^2)} \sim \mathcal{LN}(\mu + \log a, \sigma^2)$$
$$E[\exp(aX_{\mathcal{N}(\mu,\sigma^2)})] = \exp(a\mu + \tfrac{1}{2}(a\sigma)^2).$$

6.3 Chi-Squared Distribution

- $\chi^2(\nu)$ is the distribution of the sum of the squares of ν independant standard normal random variables, ie. $\sum_i Z_i^2$, $Z_i \sim \mathcal{N}(0,1)$.

-

Mean	ν
Variance	2ν
PDF	$\frac{1}{\Gamma(\nu/2)2^{\nu/2}} \exp(-x/2)\mathbf{1}_{x>0}$
Cumulants	$\nu 2^{n-1}(n-1)!$

6.4 Gamma Distribution

- The gamma distribution $\text{Gamma}(\alpha, \beta)$ (or $\text{Gamma}(k, \theta)$ on wikipedia EN) is the generalization of
  - The exponential distribution: $\text{Gamma}(1, 1/\lambda) = \text{Exp}(\lambda)$;
  - The chi-squared distribution: $\text{Gamma}(\nu/2, 2) = \chi^2(\nu)$.
- If $X_1, \dots, X_n$ are independent and $X_i \sim \text{Gamma}(\alpha_i, \beta)$, then

$$\sum_i X_i \sim \text{Gamma}(\sum_i \alpha_i, \beta).$$

-

Mean	$\alpha\beta$
Variance	$\alpha\beta^2$
Skewness	$2/\sqrt{\alpha}$
Kurtosis	$6/\alpha + 3$
PDF	$\frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp(-x/\beta)\mathbf{1}_{x>0}$

7 Estimation Theory

7.1 Maximum Likelihood Estimation

- The **maximum likelihood estimation** for a parameter $\theta \in \mathscr{R}^p$ converges to a multivariate gaussian distribution:

$$\sqrt{n}(\hat{\theta}_n - \theta) \rightsquigarrow \mathscr{N}(0, \Sigma),$$

  where Σ is the **Penrose inverse** of $\hat{I}_\theta$, the **Fisher information matrix**. In pratice, we can approximate $\hat{I}_\theta = n^{-1}\hat{H}_n$ the **hessian** at $\hat{\theta}_n$ so that

$$\Sigma = n\hat{H}_n^{-1}.$$

- The **Slutzky theorem** states that if

$$\sqrt{n}(\hat{\theta}_n - \theta) \rightsquigarrow \mathscr{N}_m(0, \Sigma),$$

  and $f : \mathscr{R}^m \to \mathscr{R}^p$, then

$$\sqrt{n}(f(\hat{\theta}_n) - f(\theta)) \rightsquigarrow \mathscr{N}_p(0, Df(\theta)\,\Sigma\,Df(\theta)^T),$$

  where $Df(\theta) \in \mathscr{R}^{p\times m}$ is the **jacobian** of f. In pratice, $Df(\theta)$ is estimated by $Df(\hat{\theta}_n)$.

8 Copulas and Multivariate Dependence

- The **margin** of a multivariate distribution is the distribution of a single random variables when all others are known to be true.

- In the bivariate case, if $(X_1, X_2) \sim F$ then F is the **joint distribution** and

$$F_1(x_1) = F(x_1, \infty),$$
$$F_2(x_2) = F(\infty, x_2)$$

are the **marginal distributions**. They are such that $X_1 \sim F_1$ and $X_2 \sim F_2$. [Remillard, Sec. 8.3.2 p. 269]

- The **rank** of an observation (x_i, y_i) out of sample $\{(x_1, y_1), \dots, (x_N, y_N)\}$ is the pair (r_{x_i}, r_{y_i}) where r_{x_i} is the rank among x_j. The **normalized rank** is the pair $(r_{x_i}, r_{y_i})/(N+1)$. [Remillard, Sec. 8.3.3 p. 269]

- Theorem. If F_1 and F_2 are two distributions, F_1 being symmetric, and the variance of $X_1 \sim F_1$ and $X_2 \sim F_2$ exist, then $X_1' = F_1^{-1}(U_1) \sim F_1$, $X_2' = F_2^{-1}(U_2) \sim F_2$ and $\mathrm{Corr}(X_1', X_2') = 0$. [Remillard, Prop. 8.4.1, p.272]

- The **Rosenblatt** transform of a copula $c(u, v)$ is given by

$$\psi(u, v) = \left(u, \frac{\partial_u c(u, v)}{\partial_u c(u, 1)}\right).$$

Therefore, if $(U, V) \sim C$, then the following identity holds

$$\psi(U, V) \sim C_\perp.$$

- For any **elliptical copula**, the relation between the **correlation matrix** ρ and the **Kendall tau** is given by

$$\tau = \frac{2}{\pi} \arcsin(\rho).$$

In particular, in the bivariate case, if $\tau = 1/2$, then $\rho = 2^{-1/2}$. [p.298]

- For **bivariate archimedean copulas**, the relation between the **generator** $\phi(t)$ and the **Kendall tau** is given by

$$\tau = 1 + 4 \int_0^1 \frac{\phi(t)}{\phi'(t)}\, dt.$$

- The **archimedean generator** $\phi(t)$ is such that

$$c(u, v) = \phi^{-1}(\phi(u) + \phi(v)).$$

8.1 Copula Families

8.1.1 The Frank Copula [p. 306]

- The **Frank copula** for $d = 2$ is given by

$$C(u, v) = \frac{1}{\log \theta} \log\left(\frac{\theta + \theta^{u+v} - \theta^u - \theta^v}{\theta - 1}\right).$$

- With $d = 2$, the following identities hold:

$$C_1 = C_\perp$$
$$C_0 = C_+$$
$$C_\infty = C_-$$

8.1.2 The Clayton Copula [p. 304]

- In the bivariate case, the **Clayton copula** is given by

$$
\begin{aligned}
C(u,v) &= (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, \qquad && 0 < \theta \\
C(u,v) &= (\max(0, u^{-\theta} + v{-}\theta - 1))^{-1/\theta}, \qquad && -1/2 \leq \theta < 0.
\end{aligned}
$$

- For $d = 2$, the **generator** is given by

$$
\phi(t) = \frac{t^{-\theta} - 1}{\theta},
$$

  for $\theta > 2$.

- With $d = 2$, the following identites hold:

$$
\begin{aligned}
C_0 &= C_\perp \\
C_\infty &= C_+ \\
C_{-1} &= C_-.
\end{aligned}
$$

- The relation between θ and τ for $d = 2$ is given by

$$
\begin{aligned}
\tau &= \frac{\theta}{\theta + 2} \\
\theta &= \frac{2\tau}{1 - \tau}.
\end{aligned}
$$

- This copula family is only **valid** for

$$
\theta \geq -\frac{1}{d-1}.
$$

8.1.3 The Gumbel Copula [p. 305]

- In the bivariate case, for $0 < \theta \leq 1$, the **Gumbel copula** is given by

$$
C(u,v) = \exp(-((-\log u)^{1/\theta} + (-\log v)^{1/\theta})^\theta)
$$

- Its generator is given by

$$
\phi(t) = (-\log t)^{1/\theta}.
$$

- The following identities hold:

$$
\begin{aligned}
C_1 &= C_\perp \\
C_0 &= C_+.
\end{aligned}
$$

- The relation between θ and τ is given by

$$
\tau = 1 - \theta.
$$

9 Calculus

- The **chain rule** of $h(x) = g(f(x))$, where $f : \mathscr{R}^n \to \mathscr{R}$ and $g : \mathscr{R} \to \mathscr{R}$ is given by

$$\nabla h(x) = g'(f(x))\nabla f(x).$$

- The infinite **Taylor expansion** of $f : \mathscr{R} \to \mathscr{R}$ around x_0 is given by

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \tfrac{1}{2}f''(x_0)(x - x_0)^2 + \tfrac{1}{3!}f'''(x_0)(x - x_0)^3 + \cdots .$$

10 Linear Algebra

- The inverse of a 2×2 matrix is given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

- The **Cholesky decomposition** of the semi-definite matrix M is the **upper triangular** matrix L such that

$$M = L^T L$$
$$L = \operatorname{chol} M.$$

Part IV
Finance Theory

11 Black-Scholes Model

- Under the **Black-Scholes Model**, the evolution of d assets $S_t \in \mathscr{R}^d$ is given elementwise by
$$S_t = S_0 \exp((\mu - \tfrac{1}{2}\sigma^2)t + \sigma W_t),$$
with $\mu, \sigma \in \mathscr{R}^d$ and where W is a d-brownian process with **correlation matrix** R, or, equivalently, covariance matrix
$$\Sigma = \operatorname{diag}(\sigma) R \operatorname{diag}(\sigma).$$
- With $b = \operatorname{chol} R$ and $\tilde{W}$ an **uncorrelated** d-brownian process, we have
$$S_t = S_0 \exp((\mu - \tfrac{1}{2}\sigma^2)t + \operatorname{diag}(\sigma) b^T \tilde{W}_t).$$
- Equivalently, with $a = \operatorname{chol} \Sigma = b \operatorname{diag} \sigma$,
$$S_t = S_0 \exp((\mu - \tfrac{1}{2}\sigma^2)t + a^T \tilde{W}_t).$$
- Matlab code, with $n = h^{-1} = 252$ looks like

```
a = chol(Sigma);
Z = randn(n,d);
X = h*ones(n,1)*(mu - 0.5*vol.^2)' - sqrt(h)*Z*a;
S = 100*exp(cumsum(X)); % nxd matrix
```

- In the bivariate case, the following relations hold:
$$a = \begin{pmatrix} \sigma_1 & \rho\sigma_2 \\ 0 & \sqrt{1-\rho^2}\sigma^2 \end{pmatrix}$$
$$\sigma_1 = a_{11}$$
$$\sigma_2 = \sqrt{a_{22}^2 + a_{12}^2}$$
$$\rho = \frac{a_{12}}{\sqrt{a_{22}^2 + a_{12}^2}}$$

11.1 Derivatives

- The value c of a regular **european call** of an asset S_t with strike K is given by
$$c = S_0 n(d_1) - K e^{-rT} n(d_2)$$
$$d_1 = \frac{\log(S_0/K) + rT + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}$$
$$d_2 = d_1 - \sigma\sqrt{T}.$$

- The **put-call parity** relates the value c of an european call with the value p of an **european put**:

$$c - p = S_0 - Ke^{-rT}.$$

- The **exchange option** with parameters q_1, q_2 over two assets $S_t^{(1)}, S_t^{(2)}$ has payoff

$$\Psi(S_t^{(1)}, S_t^{(2)}) = \max(q_2 S_T^{(2)} - q_1 S_T^{(1)}, 0).$$

Its value c is given by:

$$\begin{aligned} c &= q_2 S_0^{(2)}\, n(d_1) - q_1 S_0^{(1)}\, n(d_2) \\ \sigma^2 &= \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2 \\ d_1 &= \frac{\log(q_2 S_0^{(2)} - \log(q_1 S_0^{(1)}) + \frac{1}{2}T\sigma^2}{\sigma\sqrt{T}} \\ d_1 &= \frac{\log(q_2 S_0^{(2)} - \log(q_1 S_0^{(1)}) - \frac{1}{2}T\sigma^2}{\sigma\sqrt{T}}, \end{aligned}$$

and its delta by

$$\begin{aligned} \Delta_1 &= q_2\, n(d1) \\ \Delta_2 &= -q_1\, n(d2). \end{aligned}$$

12 Interest Rates

12.1 Ornstein Uhlenbeck Process Model

- The **Ornstein-Uhlenbeck process** model of the **spot rate** r_t is given by:

$$dr_t = \alpha(\beta - r_t)dt + \sigma\, dW_t.$$

- The Ornstein-Uhlenbeck process is invariant to new measures $\tilde{r}(t) = \lambda r(mt)$. We have the following transformation rules:

$$\begin{aligned} \tilde{\alpha} &= m\alpha \\ \tilde{\beta} &= \lambda\beta \\ \tilde{\sigma} &= \sqrt{m}\lambda\sigma \\ \tilde{q}_1 &= \sqrt{m}q_1 \\ \tilde{q}_2 &= \sqrt{m}\lambda^{-1}q_2 \end{aligned}$$

- In the **risk neutral measure**, if r_t is also an Ornstein-Uhlenbeck process, and the **market price of risk** is given by

$$q = q_1 + q_2 r,$$

then

$$d\tilde{r}_t = a(b - \tilde{r}_t)dt + \sigma\, d\tilde{W}_t.$$

- The transformation between the two measures is given by

$$a = \alpha + q_2\sigma, \qquad b = \frac{\alpha\beta - q_1\sigma}{a}.$$

- Under the **number convention**, the 1$ face value **zero-coupon bond valuation** under the risk neutral measure is given by

$$P_T = \exp(A_T - r_0\, B_T).$$

Under the **percentage convention**, it is given by

$$P_T = \exp(\tilde{A}_T - r_0\tilde{B}_T/100),$$

where

$$\begin{aligned}
B_T &= \frac{1-\exp(-aT)}{a},\\
\tilde{B}_T &= \frac{1-\exp(-aT)}{a},\\
A_T &= \left(\frac{\sigma^2}{2a^2} - b\right)(T - B_T) - \frac{\sigma^2}{4a}B_T^2,\\
\tilde{A}_T &= \left(\frac{\sigma^2}{20\,000a^2} - \frac{b}{100}\right)(T - \tilde{B}_T) - \frac{\sigma^2}{40\,000a}\tilde{B}_T^2.
\end{aligned}$$

- Under the **percentage convention/yearly basis**, The **long term yield** is given by

$$R_\infty = b - \frac{\sigma^2}{200a^2}.$$

- Under the **percentage convention/yearly basis**, the **relation between the spot rate and the yield** is given by

$$\begin{aligned}
r_0 &= \frac{T\,R_T + 100\tilde{A}_T}{\tilde{B}_T},\\
R_T &= \frac{r_0\tilde{B}_T - 100\tilde{A}_T}{T}.
\end{aligned}$$

- The Ornstein-Uhlenbeck is a **markovian, stationary and gaussian process**. Its law is given by

$$r_t \sim \mathscr{N}(\mu_t, \sigma_t^2)$$

with

$$\begin{aligned}
\mu_t &= \beta + e^{-\alpha t}(r_0 - \beta)\\
\sigma_t^2 &= \sigma^2\frac{1-e^{-2\alpha t}}{2\alpha}.
\end{aligned}$$

Furthermore, if $\alpha > 0$, then the **stationnary distribution** of r_t is

$$r_\infty \sim \mathcal{N}(\beta, \sigma^2/(2\alpha)).$$

Finally,

$$\mathrm{Cov}(r_t, r_s) = \sigma^2 \frac{e^{-\alpha(t-s)} - e^{-\alpha(t+s)}}{2\alpha}, \quad t > s$$
$$\mathrm{Cov}(r_\infty, r_t) = 0.$$

12.2 Cox-Ingersoll-Ross Model

- The **CIR model** of the spot rate r_t is given by:

$$dr_t = \alpha(\beta - r_t)dt + \sigma\sqrt{r_t}dW_t.$$

- The Feller process process is invariant to updated measures $\tilde{r}(t) = \lambda r(mt)$. We have the following transformation rules:

$$\tilde{\alpha} = m\alpha$$
$$\tilde{\beta} = \lambda\beta$$
$$\tilde{\sigma} = \sqrt{m\lambda}\sigma$$
$$\tilde{q}_1 = \sqrt{\lambda}q_1$$
$$\tilde{q}_2 = \lambda^{-1/2}q_2$$
$$\tilde{q}_2 = \lambda^{-1}q_2$$

The transformation with $m \neq 1$ doesn't carry to q.

- In the **risk neutral measure** with **market price of risk** given by

$$q = q_1 r^{-1/2} + q_2 r^{1/2}$$

then

$$d\tilde{r}_t = a(b - \tilde{r}_t)dt + \sigma\sqrt{\tilde{r}_t}d\tilde{W}_t.$$

- The transformation between the two measures is given by

$$a = \alpha + q_2\sigma, \qquad b = \frac{\alpha\beta - q_1\sigma}{a}.$$

- Under the **number convention**, the 1$ face value **zero-coupon bond valuation** under the risk neutral measure is given by

$$P_T = \exp(A_T - r_0 B_T),$$

whereas under the **percentage convention**, it is given by

$$P_T = \exp(\tilde{A}_T - r_0\tilde{B}_T/100),$$

where

$$
\begin{aligned}
\gamma &= \sqrt{a^2 + 2\sigma^2} \\
\tilde{\gamma} &= \sqrt{a^2 + 2\sigma^2/100} \\
\eta &= (\gamma + a)(1 - e^{-\gamma T}) + 2\gamma e^{-\gamma T} \\
\tilde{\eta} &= (\tilde{\gamma} + a)(1 - e^{-\tilde{\gamma} T}) + 2\tilde{\gamma} e^{-\tilde{\gamma} T} \\
B_T &= \frac{2(1 - e^{-\gamma T})}{\eta} \\
\tilde{B}_T &= \frac{2(1 - e^{-\tilde{\gamma} T})}{\tilde{\eta}} \\
A_T &= \frac{2ab}{\sigma^2} \log\left(\frac{2\gamma e^{T(a-\gamma)/2}}{\eta}\right) \\
\tilde{A}_T &= \frac{2ab}{\sigma^2} \log\left(\frac{2\tilde{\gamma} e^{T(a-\tilde{\gamma})/2}}{\tilde{\eta}}\right)
\end{aligned}
$$

- Under the **percentage convention/yearly basis**, the **long term yield** is given by

$$
R_\infty = \frac{2ab}{a + \tilde{\gamma}}.
$$

- Under the **percentage convention/yearly basis**, the **relation between the spot rate and the yield** is given by

$$
\begin{aligned}
r_0 &= \frac{T\,R_T + 100\tilde{A}_T}{\tilde{B}_T}, \\
R_T &= \frac{r_0 \tilde{B}_T - 100\tilde{A}_T}{T}.
\end{aligned}
$$

13 Lévy Models

13.1 Merton Model

- The **Merton model** has the following form:

$$
at + \sigma W_t + \sum_{i=1}^{N_t} \xi_j,
$$

with N_t a λ-Poisson process, $\xi_j \sim \mathcal{N}(\gamma, \delta^2)$ and

$$
\begin{aligned}
a &= \mu - \lambda\kappa - \tfrac{1}{2}\sigma^2, \\
\kappa &= \exp(\gamma + \delta^2/2) - 1.
\end{aligned}
$$

- Under the **risk neutral measure**, the following transformation take place

$$
\begin{aligned}
K_\phi &= e^{\zeta_0} E[e^{-\xi_i/2}] \\
\tilde{\lambda} &= K_\phi \lambda \\
\tilde{\sigma} &= \sigma \\
\tilde{\gamma} &= \gamma + \zeta_1 \delta^2 \\
\tilde{\delta} &= \delta \\
\tilde{\eta}(x) &= e^{\phi(x)} K_\phi^{-1} \eta(x),
\end{aligned}
$$

with $\eta(x)$ the density of the jumps $\xi_1 \sim \mathscr{N}(\gamma, \delta^2)$, ie.

$$
\eta(x) = \frac{1}{\delta\sqrt{2\pi}} \exp\left(-\frac{(x-\gamma)^2}{2\delta^2}\right).
$$

13.2 Kou Model

- The **Kou Model** has the following paramaters:

$$
\begin{aligned}
a &= \mu - \lambda\kappa - \tfrac{1}{2}\sigma^2, \\
\kappa &= \frac{1 + p\eta_2 + (p-1)\eta_1}{(\eta_1 - 1)(\eta_2 + 1)}, \\
\eta(x) &= p\eta_1 e^{-\eta_1 x} \mathbf{1}_{x>0} + (1-p)\eta_2 e^{-\eta_2 x} \mathbf{1}_{x\le 0}.
\end{aligned}
$$

- Under the **weighted-symmetric representation**, we have

$$
\begin{aligned}
\zeta &= \delta Y_{\mathrm{Exp}(1)} \\
\delta &= \eta_2^{-1} \\
\theta &= 0 \\
\omega &= \eta_2/\eta_1.
\end{aligned}
$$

13.3 Jump Diffusion Models

- For **jump diffusion models** of the form

$$
X_t = at + \sigma W_t + \sum_{j=1}^{N_t} \xi_j,
$$

the following identities hold

$$
\begin{aligned}
\nu &= \lambda\eta \\
\kappa = E[\exp(\xi_1) - 1] &= \int (e^x - 1)\nu(dx) \\
a &= \mu - \lambda\kappa - \tfrac{1}{2}\sigma^2.
\end{aligned}
$$

- The jump diffusion processes are always represented using their **natural characteristics**.
- The κ parameter is interpreted as the average jump size, λ is the average frequency of the jumps.
- The **martingale measure** parameters of jump diffusion processes obtained with $U_{b,\phi}$ (p. 200) are given by

$$
\begin{aligned}
K_\phi &= E[\exp(\phi(\xi_1))] \\
\tilde{\lambda} &= \lambda K_\phi \\
\tilde{\nu} &= \tilde{\lambda}\tilde{\eta} \\
\tilde{\eta} &= e^{\phi(x)} K_\phi^{-1} \eta(x) \\
\tilde{\sigma} &= \sigma \\
\tilde{\kappa} = E[\exp(\tilde{\xi}_1) - 1] &= \int (e^x - 1)\tilde{\nu}(dx) \\
\tilde{a} &= a + b\sigma + \int_{-1}^{1} x(e^{\phi(x)} - 1)\nu(dx) \\
\tilde{a}^{(\mathrm{nat})} &= a^{(\mathrm{nat})} + b\sigma
\end{aligned}
$$

- The following relation also holds:

$$
\begin{aligned}
r &= a + b\sigma + \tfrac{1}{2}\sigma^2 + \int (e^x - 1)\tilde{\nu}(dx) \\
&= \mu + b\sigma - \lambda\kappa + \tilde{\lambda}\tilde{\kappa}
\end{aligned}
$$

14 Stochastic Volatility Models

14.1 Discrete Models

- The considered models have the form

$$
\begin{aligned}
X_t &= \mu_t + \sigma_t e_t \\
&= \mu_t + \sqrt{h_t} e_t
\end{aligned}
$$

where

$$
\begin{aligned}
E[e_t|\mathscr{F}_{t-1}] &= 0 \\
\mathrm{Var}[e_t|\mathscr{F}_{t-1}] &= 1,
\end{aligned}
$$

and we have the correspondance

$$
\sigma_t^2 = h_t.
$$

- The **standard GARCH**$(1,1)$ has the form

$$
h_t = \omega + \alpha h_{t-1} e_{t-1}^2 + \beta h_{t-1}.
$$