liv7.tex **Week 7: 12.4.2016**

§7. More on European Options

1. *Bounds.*

We use the notation above. We also write c, p for the values of European calls and puts, C, P for the values of the American counterparts.

Obvious upper bounds are $c \leq S, C \leq S$, where S is the stock price (we can buy for S on the market without worrying about options, so would not pay more than this for the right to buy). For puts, one has correspondingly the obvious upper bounds $p \leq K, P \leq K$, where K is the strike price: one would not pay more than K for the right to sell at price K, as one would not spend more than one's maximum return. For lower bounds:
$c_0 \geq \max(S_0 - Ke^{-rT}, 0)$.
Proof. Consider the following two portfolios:
I: one European call plus Ke^{-rT} in cash; II: one share. Show "I $\geq$ II".
$p_0 \geq \max(Ke^{-rT} - S_0, 0)$.
Proof. By above and put-call parity.

2. *Dependence of the Black-Scholes price on the parameters.*

Recall the Black-Scholes formulae for the values c_t, p_t for the European call and put: with

$$d_{\pm} := [\log(S_t/K) + (r \pm \frac{1}{2}\sigma^2)(T-t)]/\sigma\sqrt{(T-t)},$$

$$c_t = S_t\Phi(d_+) - Ke^{-r(T-t)}\Phi(d_-), \qquad p_t = Ke^{-r(T-t)}\Phi(-d_-) - S_t\Phi(-d_+),$$

1. S. As the stock price S increases, the call option becomes more and more likely to be exercised. In the limit for large S, $d_{\pm} \to \infty$, $\Phi(d_{\pm}) \to 1$, so $c_t \to S_t - Ke^{-r(T-t)}$. This limit has a natural economic interpretation: it is the value of a *forward contract* with *delivery price* K (see e.g. Hull [H1] Ch. 3, [H2] Ch. 3).
2. σ. When the volatility $\sigma \to 0$, the stock becomes riskless, and behaves like money in the bank. Again, $d_{\pm} \to \infty$, the Black-Scholes price has the limit above, and one has the correct economic interpretation.

3. *Volatility.*

As in IV.6.6 Week 6, the volatility σ can be estimated in two ways:
a. Directly from the movement of a stock price in time [as the mathematics here is continuous time, we defer it to Ch. VI], giving what is called the *historic volatility*.

b. From the observed market prices of options: if we know everything in the Black-Scholes formula (including the price at which the option is traded) except the volatility σ, we can solve for σ. This is called *implied volatility.* Since σ appears inside the argument of the normal distribution function Φ as well as outside it, this is a transcendental equation for σ and has to be solved numerically by iteration (Newton-Raphson method). We quote (see 'The Greeks' below, and Problems 7) that the Black-Scholes price is a monotone (increasing) function of the volatility (more volatility doesn't make us 'more likely to win', but when we do win, we 'win bigger'), so there is a unique root of the equation.

In practice, one sees discrepancies between historic and implied volatility, which show limitations to the accuracy of the Black-Scholes model. But it is the standard 'benchmark model', and useful as a first approximation.

The classical view of volatility is that it is caused by future uncertainty, and shows the market's reaction to the stream of new information. However, studies taking into account periods when the markets are open and closed [there are only about 250 trading days in the year] have shown that the volatility is less when markets are closed than when they are open. This suggests that *trading itself is one of the main causes of volatility.*

Note. This observation has deep implications for the macro-prudential and regulatory issues discussed in Ch. 1. The real economy cannot afford too much volatility. Volatility is (at least partly) caused by trading. Conclusion: there is too much trading. Policy question: how can we reduce the volume of trading (much of it speculative, designed to enrich traders, and not serving a more widely useful economic purpose)? One answer is the so-called *Tobin tax* (also known as the "Robin Hood tax") (James Tobin (1918-2002), American economist; Nobel Prize for Economics, 1981). This would levy a small charge (e.g. 0.01%) on *all* financial transactions. This would both provide a major and useful source of tax revenue, and – more importantly – would discourage a lot of speculative trading, thereby (shrinking the size of the financial services industry, but) diminishing volatility, to the benefit of the general economy (Problems 7 again).

4. *The Greeks.*

These are the partial derivatives of the option price with respect to the input parameters. They have the interpretation of *sensitivities.*

(i) For a call, say, $\partial c/\partial S$ is called the *delta*, Δ. Adjusting our holdings of stock to eliminate our portfolio's dependence on S is called *delta-hedging.*

(ii) Second-order effects involve *gamma* $:= \partial(\Delta)/\partial S$.

(iii) Time-dependence is given by *Theta* is $\partial c/\partial t$.
(iv) Volatility dependence is given by *vega* $:= \partial c/\partial \sigma$.[1]

From the Black-Scholes formula (which gives the price explicitly as a function of σ), one can check by calculus (Problems 7) that

$$\partial c/\partial \sigma > 0,$$

and similarly for puts (or, use the result for calls and put-call parity). In sum: *options like volatility.* This fits our intuition. The more uncertain things are (the higher the volatility), the more valuable protection against adversity – or insurance against it – becomes (the higher the option price).
(v) *rho* is $\partial c/\partial r$, the sensitivity to interest rates.

§8. American Options.

We now consider an American call option (value C), in the simplest case of a stock paying no dividends. The following result goes back (at least) to R. C. MERTON in 1973.

Theorem (Merton's theorem). It is never optimal to exercise an American call option early. That is, the American call option is equivalent to the European call, so has the same value:

$$C = c.$$

First Proof. Consider the following two portfolios:
I: one American call option plus cash Ke^{-rT}; II: one share.
The value of the cash in I is K at time T, $Ke^{-r(T-t)}$ at time t. If the call option is exercised early at $t < T$, the value of Portfolio I is then $S_t - K$ from the call, $Ke^{-r(T-t)}$ from the cash, total

$$S_t - K + Ke^{-r(T-t)}.$$

Since $r > 0$ and $t < T$, this is $< S_t$, the value of Portfolio II at t. So Portfolio I is *always* worth less than Portfolio II if exercised *early.*

If however the option is exercised instead at expiry, T, the American call option is then the same as a European call option. We are then in

[1] Of course, vega is not a letter of the Greek alphabet! (it is the Spanish word for 'meadow', as in Las Vegas) – presumably so named for "v for volatility, variance and vega", and because vega sounds quite like beta, etc.

the situation of §7.1 above: at time T, Portfolio I is worth $\max(S_T, K)$ and Portfolio II is worth S_T. So:

$$\begin{array}{rl} \text{before} \quad T, & I < II, \\ \text{at} \quad T, & I \geq II \text{ always, and } > \text{ sometimes.} \end{array}$$

This direct comparison with the underlying [the share in Portfolio II] shows that early exercise is never optimal. Since an American option at expiry is the same as a European one, this completes the proof. //
Second Proof. One can prove the result without arbitrage arguments by using the bounds of IV.7.1. For details, see e.g. [BK, Th. 4.7.1].

Financial Interpretation.

There are two reasons why an American call should not be exercised early:
1. *Insurance.* Consider an investor choosing to hold a call option instead of the underlying stock. He does not care if the share price falls below the strike price (as he can then just discard his option) – but if he held the stock, he would. Thus the option insures the investor against such a fall in stock price, and if he exercises early, he loses this insurance.
2. *Interest on the strike price.* When the holder exercises the option, he buys the stock and pays the strike price, K. Early exercise at $t < T$ loses the interest on K between times t and T: the later he pays out K, the better.
Economic Note. Despite Merton's theorem, and the interpretation above, there are plenty of real-life situations where early exercise of an American call might be sensible, and indeed done routinely. Consider, for example, a manufacturer of electrical goods, in bulk. He needs a regular supply of large amounts of copper. The danger is future price increases; the obvious precaution is to hedge against this by buying call options. If the expiry is a year but copper stocks are running low after six months, he would exercise his American call early, to keep an adequate inventory of copper, his crucial raw material. This ensures that his main business activity – manufacturing – can continue unobstructed. Neither of the reasons above applies here:
Insurance. He doesn't care if the price of copper falls: he isn't going to sell his copper stocks, but use them.
Interest. He doesn't care about losing interest on cash over the remaining six months. He is in manufacturing to use his money to make things, and then sell them, rather than put it in the bank.

This neatly illustrates the contrast between *finance* (money, options etc.)

and *economics* (the real economy – goods and services).

Put-Call Symmetry.

The BS formulae for puts and calls resemble each other, with stock price S and discounted strike K interchanged. Results of this type are called *put-call symmetry.*

American Puts.

Recall the put-call parity of Ch. I (valid only for European options):

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

A partial analogue for American options is given by the inequalities below:

$$S - K < C - P < S - Ke^{-rT}.$$

For proof (as above) and background, see e.g. Ch. 8 (p. 216) of [H1].

We now consider how to evaluate an American put option, European and American call options having been treated already. First, we will need to work in discrete time. We do this by dividing the time-interval $[0, T]$ into N equal subintervals of length Δt say. Next, we take the values of the underlying stock to be discrete: we use the binomial model of IV.5, with a slight change of notation: we write u, d ('up', 'down') for $(1 + b), (1 + a)$: thus stock with initial value S is worth $Su^i d^j$ after i steps up and j steps down. Consequently, after N steps, there are $N + 1$ possible prices, $Su^i d^{N-i}$ $(i = 0, \cdots, N)$. It is convenient to display the possible paths followed by the stock price as a binomial tree [draw a diagram], with time going left to right and two paths, up and down, leaving each node in the tree, until we reach the $N + 1$ terminal nodes at expiry. There are 2^N possible paths through the tree. It is common to take N of the order of 30, for two reasons:
(i) typical lengths of time to expiry are measured in months (9 months, say); this gives a time-step around the corresponding number of days,
(ii) 2^{30} paths is about the order of magnitude that can be easily handled by computers (recall that $2^{10} = 1,024$, so 2^{30} is somewhat over a billion).

We now return to our treatment of the binomial model in IV.5,6, with a slight change of notation. Recall that in IV.5 (discrete time) we used $1 + r$ for the discount factor. It is convenient to call this $1 + \rho$ instead, freeing r for its usual use as the short rate of interest in continuous time. Thus $1 + \rho = e^{r\Delta t}$, and the risk-neutrality condition $p^* = (b - r)/(b - a)$ of IV.5 becomes

$$p^* = (u - e^{r\Delta t})/(u - d).$$

Now recall (IV.7) $(1+a)/(1+r) = \exp(-\sigma/\sqrt{N})$, $(1+b)/(1+r) = \exp(\sigma/\sqrt{N})$. We replaced σ^2 by $\sigma^2 T$ (to make σ the volatility per unit time), and $T = N.\Delta t$, so $\sigma/\sqrt{N}$ becomes $\sigma\sqrt{T}/\sqrt{N} = \sigma\sqrt{\Delta t}$. So now

$$u/e^{r\Delta t} = e^{\sigma/\sqrt{\Delta t}}, \qquad d/e^{r\Delta t} = e^{-\sigma\sqrt{\Delta t}}.$$

Thus $ud = e^{2r\Delta t}$. Since $\sqrt{\Delta t}$ is small, its square Δt is a second-order term; to first order, we thus have $ud = 1$, which simplifies filling in the terminal values in the binary tree.

With an eye on this simplification, we begin again: define our up and down factors u, d so that

$$ud = 1;$$

define the risk-neutral probability p^* so as to have

$$p^* = (u - e^{r\Delta t})/(u - d)$$

(to get the mean return from the risky stock the same as that from the riskless bank account), and the volatility σ to get the variance of the stock price S' after one time-step when it is worth S initially as $S^2\sigma^2\Delta t$:

$$S^2\sigma^2\Delta t = p^* S^2 u^2 + (1 - p^*)S^2 d^2 - S^2[p^* u + (1 - p^*)d]^2$$

(using $var S' = E(S'^2) - [ES']^2$). Then to first order in $\sqrt{\Delta t}$ (which is all the accuracy we shall need), one can check that we have as before

$$u = \exp(\sigma\sqrt{\Delta t}), \qquad d = \exp(-\sigma\sqrt{\Delta t}).$$

We can now calculate both the value of an American put option and the optimal exercise strategy by working backwards through the tree (this method of backward recursion in time is a form of the Dynamic Programming [DP] technique (Richard Bellman (1920-84) in 1953, book, 1957), which is important in many areas of optimization and Operational Research (OR)).
1. Draw a binary tree showing the initial stock value and having the right number, N, of time-intervals.
2. Fill in the stock prices: after one time interval, these are Su (upper) and Sd (lower); after two time-intervals, Su^2, S and $Sd^2 = S/u^2$; after i time-intervals, these are $Su^j d^{i-j} = Su^{2j-i}$ at the node with j 'up' steps and $i - j$ 'down' steps (the '(i, j)' node).
3. Using the strike price K and the prices at the *terminal nodes*, fill in the

payoffs ($f_{N,j} = \max[K - Su^j d^{N-j}, 0]$) from the option at the terminal nodes (where, at expiry, the values of the European and American options coincide) underneath the terminal prices.

4. Work back down the tree one time-step. Fill in the 'European' value at the penultimate nodes as the discounted values of the upper and lower right (terminal node) values, under the risk-neutral measure – 'p^* times lower right plus $1 - p^*$ times upper right' [notation of IV.6 Week 6]. Fill in the 'intrinsic' (or early-exercise) value - the value if the option is exercised. Fill in the American put value as the higher of these.

5. Treat these values as 'terminal node values', and fill in the values one time-step earlier by repeating Step 4 for this 'reduced tree'.

6. Iterate. The value of the American put at time 0 is the value at the root – the last node to be filled in. The 'early-exercise region' is the node set where the early-exercise value is the higher; the rest is the 'continuation region'.

Note. The above procedure is simple to describe and understand, and simple to programme. It is laborious to implement numerically by hand, on examples big enough to be non-trivial. Numerical examples are worked through in detail in [H1], 359-360 and [CR], 241-242.

Mathematically, the task remains of describing the *continuation region* – the part of the tree where early exercise is not optimal. This is a classical *optimal stopping problem*. No explicit solution is known (and presumably there isn't one). We will, however, connect the work above with that of III.7 [Week 5] on the Snell envelope. Consider the pricing of an American put, strike price K, expiry N, in discrete time, with discount factor $1+r$ per unit time as earlier. Let $Z = (Z_n)_{n=0}^N$ be the payoff on exercising at time n. We want to price Z_n, by U_n say (to conform to our earlier notation), so as to avoid arbitrage; again, we work backwards in time. The recursive step is

$$U_{n-1} = \max(Z_{n-1}, \frac{1}{1+r} E^*[U_n | \mathcal{F}_{n-1}]),$$

the first alternative on the right corresponding to early exercise, the second to the discounted expectation under P^*, as usual. Let $\tilde{U}_n = U_n/(1+r)^n$ be the discounted price of the American option. Then

$$\tilde{U}_{n-1} = \max(\tilde{Z}_{n-1}, E^*[\tilde{U}_n | \mathcal{F}_{n-1}]) :$$

$(\tilde{U}_n)$ is the *Snell envelope* (III.7) of the discounted payoff process $(\tilde{Z}_n)$, so:
(i) a P^*-supermartingale,

(ii) the smallest supermartingale dominating $(\tilde{Z}_n)$,
(iii) the solution of the optimal stopping problem for $\tilde{Z}$.
P-measure and P^*- (or $Q-$) measure.

We use P and P^* in the above, as E and E^* are convenient, but P and Q when the emphasis is on Q, for brevity.

The measure P, the *real* (or real-world) probability measure, models the uncertainty driving prices, which are indeed uncertain, thus allowing us to bring mathematics to bear on financial problems. But P is difficult to get at directly. By contrast, Q is more accessible: the *market* tells us about Q, or more specifically, *trading* does. In addition, trading also tells us about the *volatility* σ, via implied volatility, which we can infer from observing the prices at which options are traded. So Q is certainly more accessible than P. There is thus a sense in which it is Q, rather than P, which is the more real.

It is as well to bear all this in mind when looking at specific problems, particularly numerical ones. Now that we know the CRR binomial-tree model, which gives us the *Black-Scholes formula* (in discrete time, and by the limiting argument above, in continuous time), the main result of the course, we can recognise the 'one-period, up or down' model ($/SFr in I.8 L5, price of gold in Problems 5), though clearly artificial and stylised, as a workable 'building block' of the whole theory. Because P itself does not occur in the Black-Scholes formula(e), from a purely financial point of view there is little need to try to construct more realistic, and so more complicated, models of P. Instead, one can exploit what one can infer about Q, which does occur in Black-Scholes, from seeing the prices at which options trade.

From the economic point of view, it is the real world, the real economy, and so the real probability measure P, that matters. The 'Q-measure-eye view of the world' has a degree of artificiality, in so far as options do. One can eat food, and needs to. One can't eat options.

A fuller discussion of Q-measure involves *Arrow-Debreu* prices, equilibria etc., but we omit this for lack of time.
Where we are.

The course splits neatly into three parts: Ch. I, II on background, Ch. III, IV on discrete time, and Ch. V, VI on continuous time. We have already seen the main ideas – and proved nearly everything seen so far. In V, VI we gain the tremendous power of Itô (stochastic) calculus (calculus is our most powerful weapon, in mathematics and science!), and the ability to work in continuous time. What we lose is the ability to prove so much and to see what is happening so clearly and so concretely.

Time.

Is time discrete or continuous? It is both (or from the above we have wasted two chapters!). If we work in discrete time we "have a digital watch". If we work in continuous time, we "have a watch with a sweep second hand". Continuous time is harder, *but* we can then use *calculus*, as above.

Chapter V. STOCHASTIC PROCESSES IN CONTINUOUS TIME

§1. Filtrations; Finite-Dimensional Distributions

The underlying set-up is as before, but now *time is continuous rather than discrete*; thus the time-variable will be $t \geq 0$ in place of $n = 0, 1, 2, \ldots$. The information available at time t is the σ-field $\mathcal{F}_t$; the collection of these as $t \geq 0$ varies is the filtration, modelling the information flow. The underlying probability space, endowed with this filtration, gives us the stochastic basis (filtered probability space) on which we work.

We assume that the filtration is *complete* (contains all subsets of null-sets as null-sets), and *right-continuous*: $\mathcal{F}_t = \mathcal{F}_{t+}$, i.e.

$$\mathcal{F}_t = \cap_{s>t} \mathcal{F}_s$$

(Meyer's 'usual conditions' – right-continuity and completeness).

A stochastic process $X = (X_t)_{t \geq 0}$ is a family of random variables defined on a filtered probability space with X_t $\mathcal{F}_t$-measurable for each t: thus X_t is known when $\mathcal{F}_t$ is known, at time t.

If $\{t_1, \cdots, t_n\}$ is a finite set of time-points in $[0, \infty)$, $(X_{t_1}, \cdots, X_{t_n})$, or $(X(t_1), \cdots, X(t_n))$ (for typographical convenience, we use both notations interchangeably, with or without ω: $X_t(\omega)$, or $X(t, \omega)$) is a random n-vector, with a distribution, $\mu(t_1, \cdots, t_n)$ say. The class of all such distributions as $\{t_1, \cdots, t_n\}$ ranges over all finite subsets of $[0, \infty)$ is called the class of all *finite-dimensional distributions* of X. These satisfy certain obvious consistency conditions:
(i) deletion of one point t_i can be obtained by 'integrating out the unwanted variable', as usual when passing from joint to marginal distributions,
(ii) permutation of the t_i permutes the arguments of $\mu(t_1, \cdots, t_n)$ on $\mathbb{R}^n$.
Conversely, a collection of finite-dimensional distributions satisfying these two consistency conditions arises from a stochastic process in this way (this is the content of the DANIELL-KOLMOGOROV Theorem: P. J. Daniell in

1918, A. N. Kolmogorov in 1933).

Important though it is as a general existence result, however, the Daniell-Kolmogorov theorem does not take us very far. It gives a stochastic process X as a random function on $[0, \infty)$, i.e. a random variable on $\mathbb{R}^{[0,\infty)}$. This is a vast and unwieldy space; we shall usually be able to confine attention to much smaller and more manageable spaces, of functions satisfying regularity conditions. The most important of these is *continuity*: we want to be able to realise $X = (X_t(\omega))_{t\geq 0}$ as a random *continuous* function, i.e. a member of $C[0, \infty)$; such a process X is called *path-continuous* (since the map $t \mapsto X_t(\omega)$ is called the sample path, or simply path, given by ω) – or more briefly, *continuous*. This is possible for the extremely important case of *Brownian motion* (below), for example, and its relatives. Sometimes we need to allow our random function $X_t(\omega)$ to have jumps. It is then customary, and convenient, to require X_t to be *right-continuous with left limits* (rcll), or càdlàg (continu à droite, limite à gauche) – i.e. to have X in the space $D[0, \infty)$ of all such functions (the *Skorohod space*). This is the case, for instance, for the *Poisson process* and its relatives.

General results on realisability – whether or not it is possible to *realise*, or obtain, a process so as to have its paths in a particular function space – are known, but it is usually better to *construct* the processes we need directly on the function space on which they naturally live.

Given a stochastic process X, it is sometimes possible to improve the regularity of its paths without changing its distribution (that is, without changing its finite-dimensional distributions). For background on results of this type (separability, measurability, versions, regularization, ...) see e.g. Doob's classic book [D].

The continuous-time theory is technically much harder than the discrete-time theory, for two reasons:
(i) questions of path-regularity arise in continuous time but not in discrete time,
(ii) *uncountable* operations (like taking sup over an interval) arise in continuous time. But measure theory is constructed using *countable* operations: uncountable operations risk losing measurability.

Filtrations and Insider Trading

Recall that a filtration models an information flow. In our context, this is the information flow on the basis of which market participants – traders, investors etc. – make their decisions, and commit their funds and effort.

All this is information in the *public* domain – necessarily, as stock exchange prices are publicly quoted.

Again necessarily, many people are involved in major business projects and decisions (an important example: mergers and acquisitions, or M&A) involving publicly quoted companies. Frequently, this involves price-sensitive information. People in this position are – rightly – prohibited by law from profiting by it directly, by trading on their own account, in publicly quoted stocks but using private information. This is rightly regarded as theft at the expense of the investing public.[2] Instead, those involved in M&A etc. should seek to benefit legitimately (and indirectly) – enhanced career prospects, commission or fees, bonuses etc.

The regulatory authorities (Securities and Exchange Commission – SEC – in US; Financial Conduct Authority (FCA) and Prudential Regulation Authority (PRA, part of the Bank of England (BoE) in UK) monitor all trading electronically. Their software alerts them to patterns of suspicious trades. The software design (necessarily secret, in view of its value to criminals) involves all the necessary elements of Mathematical Finance in exaggerated form: economic and financial insight, plus: mathematics; statistics (especially pattern recognition, data mining and machine learning); numerics and computation.

§2. Classes of Processes.

1. *Martingales.*

The martingale property in continuous time is just that suggested by the discrete-time case:

$$E[X_t|\mathcal{F}_s] = X_s \qquad (s < t),$$

and similarly for submartingales and supermartingales. There are regularization results, under which one can take X_t right-continuous in t. Among the contrasts with the discrete case, we mention that the Doob-Meyer decomposition, easy in discrete time (III.8), is a deep result in continuous time. For background, see e.g.
MEYER, P.-A. (1966): *Probabilities and potentials*. Blaisdell
– and subsequent work by Meyer and the French school (Dellacherie & Meyer, *Probabilités et potentiel*, I-V, etc.

[2]The plot of the film *Wall Street* revolves round such a case, and is based on real life – recommended!

2. Gaussian Processes.

Recall the multivariate normal distribution $N(\mu, \mathbf{\Sigma})$ in n dimensions. If $\mu \in \mathbb{R}^n$, $\mathbf{\Sigma}$ is a non-negative definite $n \times n$ matrix, $\mathbf{X}$ has distribution $N(\mu, \mathbf{\Sigma})$ if it has characteristic function

$$\phi_{\mathbf{X}}(\mathbf{t}) := E \exp\{i\mathbf{t}^T.\mathbf{X}\} = \exp\{i\mathbf{t}^T.\mu - \frac{1}{2}\mathbf{t}^T\mathbf{\Sigma}\mathbf{t}\} \qquad (\mathbf{t} \in \mathbb{R}^n).$$

If further $\mathbf{\Sigma}$ is positive definite (so non-singular), $\mathbf{X}$ has density

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{1}{2}n}|\Sigma|^{\frac{1}{2}}} \exp\{-\frac{1}{2}(\mathbf{x} - \mu)^T \mathbf{\Sigma}^{-1}(\mathbf{x} - \mu)\}$$

(*Edgeworth's Theorem* of 1893: F. Y. Edgeworth (1845-1926), English statistician).

A process $X = (X_t)_{t \geq 0}$ is *Gaussian* if all its finite-dimensional distributions are Gaussian. Such a process can be specified by:
(i) a measurable function $\mu = \mu(t)$ with $EX_t = \mu(t)$,
(ii) a non-negative definite function $\sigma(s, t)$ with

$$\sigma(s, t) = cov(X_s, X_t).$$

Gaussian processes have many interesting properties. Among these, we quote *Belayev's dichotomy*: with probability one, the paths of a Gaussian process are either continuous, or extremely pathological: for example, unbounded above and below on any time-interval, however short. Naturally, we shall confine attention in this course to continuous Gaussian processes.

3. Markov Processes.

X is *Markov* if for each t, each $A \in \sigma(X_s : s > t)$ (the 'future') and $B \in \sigma(X_s : s < t)$ (the 'past'),

$$P(A|X_t, B) = P(A|X_t).$$

That is, if you know where you are (at time t), how you got there doesn't matter so far as predicting the future is concerned – equivalently, past and future are conditionally independent given the present.

The same definition applied to Markov processes in discrete time.

X is said to be *strong Markov* if the above holds with the *fixed* time t replaced by a *stopping time* T (a random variable). This is a real restriction of the Markov property in continuous time (though not in discrete time).