

## 1 Probability theory overview

**1.1 Cumulative distribution function (cdf)** Every real-valued random variable  $X$  can be described by its cumulative distribution function (*cdf*)  $F_X : \mathbb{R} \rightarrow [0, 1]$  defined as

$$F_X(x) = \mathbb{P}[X \leq x]$$

Notice that, as mentioned above,  $\{X \leq x\} = X^{-1}((-\infty, x]) \in \mathcal{F}$ .

### Discrete random variable

We call var. discrete, if its image (pmf) is **finite** or **countably finite**.

### Continuos random variables

We call radom var continuous if its cdf is continous. It implies the image of  $X$  (pdf) is “uncountably infite”.

## 1.2 Quantile functions and quantiles

Quantile is \*\*generalized inverse of the cdf  $F_x$  of  $X$ , defined as:

$$Q_X(p) = \inf \{x \in \mathbb{R} \mid p \leq F_X(x)\}$$

For some probability level  $\alpha \epsilon (0, 1)$ , the  $\alpha$  – *quantile* is just the value  $Q_x(\alpha)$

## 1.3 Expected Value

or expectation, mean, average, first moment- characterises the tendency or long-run average value

**For absolutely continuous**

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

**For discrete**

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} x_k \mathbb{P}[X = x_k] = \sum_{k=1}^{\infty} x_k p_k$$

**Main take-home** Example. Let  $X \sim \mathcal{U}(a, b)$ , (uniformly distributed) then the expected value of  $X$  is

$$\begin{aligned} \mathbb{E}[X] &= \int_{-\infty}^{\infty} x f_X(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{x^2}{2(b-a)} \Bigg|_{x=a}^b \\ &= \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2} \end{aligned}$$

$$\mathbb{E}[X] = \frac{a+b}{2}$$

## 1.4 Variance and SD

- variance: distance to the mean, expectation of the squared deviation from its mean
- standard deviation: sqare root of variance

$$\text{Var}(X) = \mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

## 1.5 Covariance between two RV

expected value of the product of their SD.

$$\text{cov}(X, Y) = \frac{\sigma_{XY}}{N-1} = \frac{\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]}{N-1}$$

$$\text{cov}(X, Y) = \frac{\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]}{N-1}$$

## 1.6 Pearson’s correlation

Measure of linear correlation between two RV.

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sqrt{\mathbb{V}[X]}\sqrt{\mathbb{V}[Y]}} = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y}$$

**! If two vars are independent, they are also uncorrelated, the opposite must not be true.**

$$\mathbb{V}[X + Y] = \mathbb{V}[X] + \mathbb{V}[Y] + 2\text{cov}(X, Y)$$

Note that variance of a sum is sum of variance only if vars are uncorrelated.

## 1.7 Skewness

Measure of the assymetry of the distribution around its mean. Third standardized moment.

$$\gamma_1(X) = \mathbb{E} \left[ \left( \frac{X - \mathbb{E}[X]}{\sqrt{\mathbb{V}[X]}} \right)^3 \right] = \mathbb{E} \left[ \left( \frac{X - \mu}{\sigma} \right)^3 \right]$$

## 1.8 Kurtosis

Measure of the width of its tails. Fourth standadized moment. Kurtosis = 3 is average, above (below) is excess kurtosis.

$$\text{Kurt}[X] = \gamma_2(X) = \kappa(X) = \mathbb{E} \left[ \left( \frac{X - \mu}{\sigma} \right)^4 \right]$$

## 1.9 Independence

We say two elements  $A$  and  $B$  of a  $\sigma$  -algebra are independent and we write  $A \perp B$  if

$$\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$$

To see why we call this property independence, we recall the definition of the conditional probability and use ‘independence’,

$$\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} = \frac{\mathbb{P}[A]\mathbb{P}[B]}{\mathbb{P}[B]} = \mathbb{P}[A]$$

So the probability of  $A$  stays the same independent of the occurrence of  $B$ . The same is true the other way around

$$\mathbb{P}[B \mid A] = \frac{\mathbb{P}[B \cap A]}{\mathbb{P}[A]} = \mathbb{P}[B]$$

## 1.10 Normal distribution

**probability density function**

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( -\frac{(x-\mu)^2}{2\sigma^2} \right).$$

**density function , cdf**

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{x^2}{2} \right)$$

**Combined**

$$f_X(x) = \frac{1}{\sigma} \varphi \left( \frac{x-\mu}{\sigma} \right)$$

## 1.11 Log-normal distribution

We say  $X$  is log-normally distributed and write  $X \sim \log \mathcal{N}(\mu, \sigma^2)$  if

$$Y = \log(X) \sim \mathcal{N}(\mu, \sigma^2)$$

$$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp \left[ -\frac{(\log x - \mu)^2}{2\sigma^2} \right]$$

## 1.12 t-distribution

Like normal but with heavier tails. **cdf**

$$t_\nu(x) = \int_{-\infty}^x \frac{\Gamma\left(\frac{\nu+1}{2}\right) \nu^{\frac{\nu}{2}}}{\sqrt{\pi} \Gamma\left(\frac{\nu}{2}\right)} \left(\nu + t^2\right)^{-\frac{\nu+1}{2}} dt$$

Where  $\nu$  denotes the degrees of freedom.

## 2 Bonds

## 2.1 Present value of Zero coupon bonds

$$PV(y) = \frac{C_T}{(1+y)^T}$$

Notes:

- market value of zero-coupon bonds decreases with longer maturities
- keeping T fixed, the market value of the zero-coupon bond decreases as the yield increases

Semi-annual compounding (e.g. used in the U.S. Treasury bond market): interest rate  $y_s$  is derived from:

$$P(y_s) = \frac{C_T}{\left(1 + \frac{y_s}{2}\right)^{2T}}$$

Continuous compounding (used ubiquitously in the quantitative finance literature) interest rate  $y_c$  is derived from:

$$P(y_c) = \frac{C_T}{\exp(y_c T)} = e^{-y_c T} C_T$$

$$yield = \left(\frac{FV}{CV}\right)^{\frac{1}{T}} - 1$$

$$1 + y_{p.a.} = \left(1 + \frac{y_{p.s.}}{2}\right)^2 = \exp(y_{continuous})$$

**Price of a coupon paying bond**

$$P(y) = \frac{cF}{y} \left(1 - \frac{1}{(1+y)^T}\right) + \frac{F}{(1+y)^T}$$

Price of a perpetual bond/consol:  $P(y) = \frac{c}{y} F$

## 2.2 Taylor series

$$\begin{aligned} T_{f,x_0}(x) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \\ &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2} (x - x_0)^2 + \dots \end{aligned}$$

$$\sum_{n=0}^2 \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2} (x - x_0)^2$$

a Taylor polynomial of degree two and so on.

**2.3 Duration**  $\triangleright$  The Dollar-Duration ( $DD$ ) is the negative of the first derivative

$$DD(y) = -P'(y)$$

$\triangleright$  The Modified Duration ( $D^*$ ) is a normalized version of the Dollar-Duration:

$$D^*(y) = \frac{DD(y)}{P(y)} = -\frac{P'(y)}{P(y)}$$

$\triangleright$  The Macaulay Duration ( $D_M$ ), yet another normalized version:

$$D_M(y) = (1+y)D^*(y) = -(1+y)\frac{P'(y)}{P(y)}$$

Duration is average cash-flow during maturity, or average time at which cash is received. Measures sensitivity to change in yields.

**2.4 Convexity** Improved approximation of bond price sensitivity given by duration. Taylor series up to the second derivative. Hence:  $\triangleright$  Again, we take the second derivative of each term in the sum separately. We get for all  $t \in \{1, \dots, T\}$ ,

$$\frac{d^2}{dy^2} \left( \frac{C_t}{(1+y)^t} \right) = \frac{d}{dy} \left( -\frac{tC_t}{(1+y)^{t+1}} \right) = \frac{t(t+1)C_t}{(1+y)^{t+2}}$$

$\triangleright$  Hence, the second derivative of the bond price is given by

$$\begin{aligned} P''(y) &= \frac{d^2 P}{dy^2}(y) = \sum_{t=1}^T \frac{t(t+1)C_t}{(1+y)^{t+2}} = \frac{1}{(1+y)^2} \sum_{t=1}^T \frac{t(t+1)C_t}{(1+y)^t} \\ &= \frac{1}{(1+y)^2} \left( \frac{2C_1}{1+y} + \frac{6C_2}{(1+y)^2} + \dots + \frac{T(T+1)C_n}{(1+y)^T} \right). \end{aligned}$$

**Dollar convexity** ( $DC$ ) is defined as the second derivative of the bond price with respect to the yield, that means we set

$$DC(y)=\frac{d^2P}{dy^2}(y)=P''(y)$$

#### Bond convexity  
( $\kappa$ ) is defined to be the second derivative of the bond price with respect to the yield, divided by the bond price, that is

$$\kappa(y)=\frac{DC(y)}{P(y)}=\frac{P''(y)}{P(y)}$$

Note that some practitioners refer to convexity as the second derivative divided by the present value and additionally multiplied by the factor  $\frac{1}{2}$

**Zero Coupon convexity**

$$\kappa(y)=\frac{(T+1)T}{(1+y)^2}$$

**Conclusions:**

- for small movements, duration (linear approx.) is reasonably good.
- dollar duration measures the negative slope of the tanget to the price-yield curve at the point  $y_0$
- duration is the **average time to wait for each payment weighted by their present values**

### 3 Econometrics (volatility)

- vol alludes to the fact that asset price, thus the return, is **random**
- vol is a measure of uncertainty of the stock returns
- typical vol of stocks is between 20-50%
- matters for rebalancing positions

#### 3.1 Stylized facts

- returns of stocks exhibit **heteroscedasticity and positive autocorrelation**
- volatility clusters - large (small) vol is followed by also large(small) vol, of either sign
- autocorrelations are not large and decrease more or less slowly with the lag order

#### 3.2 Intro Unbiased estimator of the variance

$$\hat{v}=\frac{1}{n-1}\sum_{i=1}^n(u_i-\bar{u})^2$$

The annualized volatility  $\hat{\sigma}$ , under the assumption of Gaussianity, can be estimated as

$$\hat{\sigma}=\frac{\sqrt{\hat{v}}}{\sqrt{\tau}}$$

$$SE=\frac{\hat{\sigma}}{\sqrt{2n}}.$$

#### 3.3 ARCH(m)

- captures dependence in the form of positive autocorrelations in the square residuals
- estimate of the variance is based on LT average variance and m observations. The older it is, less weight it's given.

**Creation**

$$\sigma_n^2=\sum_{i=1}^m\alpha_iu_{n-i}^2$$

▷ The coefficient  $\alpha_i>0$  is the weight given to the observation  $i$  days ago. **Alpha serves to give higher weights to more recent observations**  
Extension is the assumption there is a long-run variance rate and that it should be given some weight.

$$\sigma_n^2=\gamma V_L+\sum_{i=1}^m\alpha_iu_{n-i}^2$$

where  $V_L$  is the long-run variance rate and  $\gamma$  is the weight assigned to  $V_L$ . Often,  $\omega=\gamma V_L$ . Because the weights must sum to unity, we have

$$\gamma+\sum_{i=1}^m\alpha_i=1.$$

**EWMA Model**

- designed to track changes in the volatility

Weights  $\alpha_i$  **decrease exponentially** as we move back through time.

$$\alpha_{i+1}=\lambda\alpha_i,\text{ where }\lambda\in[0,1]$$

It turns out that this weighting scheme leads to a particularly simple formula for updating volatility estimates:

$$\sigma_n^2=\lambda\sigma_{n-1}^2+(1-\lambda)u_{n-1}^2$$

$$\sigma_n^2=(1-\lambda)\sum_{i=1}^m\lambda^{i-1}u_{n-i}^2+\lambda^mu_{n-m}^2$$

Lambda is empirically at 0.94.

**GARCH (1,1)**

- vol. clustering often generates more extreme values and less central values compared to case when returns are independent, i.e.  $\alpha=\beta=0$
- incorporates mean-reversion, whereas EWMA does not

The conditional variance  $\sigma_t^2$  is calculated from a long-run average variance rate  $V_L$ , the past variance  $\sigma_{t-1}^2$ , and the past return  $u_{t-1}$ :

$$\sigma_t^2=\underbrace{\gamma\cdot V_L}_{\omega}+\beta\cdot\sigma_{t-1}^2+\alpha\cdot u_{t-1}^2,$$

Where:

- conditional variance  $\sigma_{t-1}^2$  is calculated from the LT average  $V_L$ , past return  $u_{t-1}$
- $1=\gamma+\beta+\alpha$

### 4 Copulas Correlation measures only **linear dependency**- copulas solve this problem.

**Definition** Copulas are functions that separate the **marginal distributions** from the **dependency structure** of a multivar distribution  
Copulas express dependence on a **quantile scale**- useful for dependence of **extreme outcomes**. They faciliate **bottom-up approach**- that is better because we often understand marginal behavior of individual risk factors than the aggregate, whole dependence structure.

**Definition 2** Coupula is a function with 2 vars and has 3 properties:

- groundedness
- margins
- increasing

#### 4.1 Grounded functions

**Definition.** We say a function  $H:X\times Y\rightarrow\mathbb{R}$  is grounded if

$$H(x,m_Y)=H(m_X,y)=0\text{ for all } (x,y)\in X\times Y$$

Example. Let  $H:[-1,1]\times[0,\infty]\rightarrow\mathbb{R}$  be defined by

$$H(x,y)=\frac{(x+1)(e^y-1)}{x+2e^y-1}$$

Then  $H$  is grounded, since  $H(x,0)=H(-1,y)=0$  for all  $x\in[-1,1]$  and  $y\in[0,\infty]$ .

#### 4.2 Margins

Suppose now that  $M_X=\max X$  and  $M_Y=\max Y$  exist.

**Definition.** The margins of a function  $H:X\times Y\rightarrow\mathbb{R}$  are the functions

$$\begin{aligned}F:X&\rightarrow\mathbb{R}\text{ defined by }F(x)=H(x,M_Y)\text{ and}\\G:Y&\rightarrow\mathbb{R}\text{ defined by }G(y)=H(M_X,y)\end{aligned}$$

#### 4.3 2-increasing functions

**Definition.** Let  $A=[x_1,x_2]\times[y_1,y_2]\subset X\times Y$ . The  $H$ -volume of the rectangle  $A$  is defined as

$$V_H(A)=H(x_2,y_2)-H(x_2,y_1)-H(x_1,y_2)+H(x_1,y_1).$$

We abbreviate this as

$$V_H(A)=\Delta_{y_1}^y\Delta_{x_1}^xH(x,y)$$

for the first order differences of  $H$  on  $A$ .  $\Delta_{x_1}^xH(x,y)=H(x_2,y)-H(x_1,y)$  and  $\Delta_{y_1}^yH(x,y)=H(x,y_2)-H(x,y_1)$ . We call a function  $H:X\times Y\rightarrow\mathbb{R}$  2-increasing 1 if  $V_H(A)\geq 0$  for all rectangles  $A\subset X\times Y$

### 4.4 Summary of copulas with properties

**Definition.** A function  $C:[0,1]^2\rightarrow[0,1]$  is called a copula (or more precisely, a two-dimensional copula, or a 2-copula) if 1.  $C$  is grounded, that means  $C(u,0)=C(0,v)=0$ , 2. Its margins are the identity, that means for every  $u,v\in[0,1]$

$$C(u,1)=u\text{ and }C(1,v)=v$$

3.  $C$  is 2 -increasing, that means for every  $u_1,u_2,v_1,v_2\in[0,1]$  with  $u_1\leq u_2,v_1\leq v_2$  we have

$$C(u_2,v_2)-C(u_2,v_1)-C(u_1,v_2)+C(u_1,v_1)\geq 0$$

### 4.5 Types of copulas

**Example.** The fundamental examples are: ▷ Product or independence copula:  $\Pi(u,v)=uv$ , ▷ Lower Fréchet-Hoeffding bound:  $W(u,v)=\max(u+v-1,0)$ , ▷ Upper Fréchet-Hoeffding bound:  $M(u,v)=\min(u,v)$ .

### 4.6 Sklar's theorem

**Theorem.** (Sklar's theorem) 1. Let  $H$  be a joint distribution function with margins  $F$  and  $G$ . Then there exists a copula  $C$  such that for all  $x,y\in\mathbb{R}$

$$H(x,y)=C(F(x),G(y))$$

If  $F$  and  $G$  are continuous, then  $C$  is unique. Otherwise,  $C$  is uniquely determined on  $\text{Ran }F\times\text{Ran }G$ , where  $\text{Ran }F$  denotes the range of  $F$ . 2. Conversely, if  $C$  is a copula and  $F$  and  $G$  are distribution functions, then the function  $H$ , given in Equation (1), defines a joint distribution function with margins  $F$  and  $G$ .

**THEN THE ASSOCIATED COPULA IS:**

$$C(u,v)=\frac{uv}{u+v-uv}$$

## Examples of copulas

- fundamental couplas: independent copula, upper and lower bounds
- implicit copulas: extracted from known cdf's F via Sklar's theorem. Cannot write  $C\equiv\dots\dots$  directly.
- explicit copulas: can write  $C\equiv\dots\dots$  directly, but their use is very limited and they are quite unnecessary.

**Parametric copulas** We can divide parametric couplas into *implicit* and *explicit* ones. Sklar's theorem states that we can always find a copula in a parametric distribution function. Let  $H$  be the distribution function and let  $F_1,F_2$  its continuous margins. Then the implied copula is

$$C(u_1,u_2)=H\left(F_1^{(-1)}(u_1),F_2^{(-1)}(u_2)\right)$$

Such a copula may not have a simple closed form.

**Financial interpretation** ▷ Consider 2 daily returns  $X=(X_1,X_2)$  with pairwise correlations all equal to  $\rho=0.5$ .

However, we are unsure about the best joint model.

$D$  If the copula of  $X$  is  $C_{\rho=0.5}^{\text{Gauss}}$ , the probability that on a given day both returns lie below their  $\alpha=0.01$  quantiles is (this is a coupla)

$$\mathbb{P}\left[X_1\leq F_1^{-1}(\alpha),X_2\leq F_2^{-1}(\alpha)\right]=\mathbb{P}[U_1\leq\alpha,U_2\leq\alpha]\approx 1.29\times 10^{-3}$$

In the long run such an event will happen approximately once every 775 trading days ( $\approx$  once every 3 years

▷ If the copula of  $X$  is  $C_{\nu=4,\rho=0.5}^t$  (student distribution) however, such an event will happen approximately 2.22 times more often, i.e.  $\approx$  once every 1.3 years.!!!!

### 5 Guest Lecture - Modelling dependencies

#### 5.1 Pearson correlation

$$\rho(X_1,X_2)=\frac{\text{Cov}(X_1,X_2)}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}}$$

where  $\text{Cov}(X_1,X_2)=E(X_1X_2)-E(X_1)E(X_2)$

- parametric
- only captures linear dependeces (no tail risk)

#### 5.2 Spearman $\rho_S(X_1,X_2)=\rho(F_1(X_1),F_2(X_2))$

- not sensitive to outliers -good for heavy tailed distributions

#### 5.3 Kendall $\tau(X_1,X_2)=E\left[\text{sign}\left((X_1-\bar{X}_1)(X_2-\bar{X}_2)\right)\right]$

where  $\bar{X}_1$  and  $\bar{X}_2$  have the same joint distribution, but are independent of  $X_1$  and  $X_2$

- not sensitive to outliers -good for heavy tailed distributions
- **!appropriate to measure tail behavior**

### 6 Discrete Time Models

### 6.1 Adaptedness and predictability

We call an indexed family of RV a **stochastic process**  
▷ A stochastic process  $Z = (Z_t)_{t=0,\dots,T}$  is called  $\mathbb{F}$ -adapted if  $Z_t$  is  $\mathcal{F}_{(J)}$  measurable for every  $t = 0, \dots, T$ .  
▷ A stochastic process  $Y = (Y_t)_{t=1,\dots,T}$  is called  $\mathbb{F}$ -predictable if  $Y_t$  is  $\mathcal{F}_{(t-x)}$ -measurable for every  $t = 1, \dots, T$ . In our model, the price process  $S = (S_t)_{t=0,\dots,T}$  is  $\mathbb{F}$ -adapted.

### 6.2 Trading strategies

$\sum \xi_t^k$  represents the number of shares of asset  $k$  held during the  $t^{\text{th}}$  trading period between times  $t-1$  and  $t$ .  $\sum \xi_t^k S_{t-1}^k$  denotes the amount invested in the  $k^{\text{th}}$  asset at time  $t-1$ , while  $\xi_t^k S_t^k$  is the resulting value at time  $t$ .  
D Predictability of the strategy represents the fact that any investment must be allocated at the beginning of each trading period, without anticipating future prices.

#### Valuation process

The (discounted) value process  $V = (V_t)_{t \in \{0,\dots,T\}}$  of a trading strategy  $\xi$  is given by  $\Rightarrow V_0 = \xi_1 \cdot \bar{X}_0$  and  $V_t = \xi_t \cdot \bar{X}_t$  for  $t = 1, \dots, T$ .  $V_t$  represents the portfolio value at the end of the  $t^{\text{th}}$  trading period.

**Arbitrage opportunities** Arbitrage is a self-financing, risk-free trading strategy such that:  $V_0 \leq 0$ ;  $V_T \geq 0$

**Martingale definition** Definition. Let  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$  be a filtered probability space. Then a stochastic process  $M = (M_t)_{t=0,\dots,T}$  is called a  $(\mathbb{P}, \mathbb{F})$ -martingale if

▷  $M$  is  $\mathbb{F}$ -adapted,  
▷  $\mathbb{E}_{\mathbb{P}}[|M_t|] < \infty$  for every  $t \in \{0, \dots, T\}$   
▷  $\mathbb{E}_{\mathbb{P}}[M_t | \mathcal{F}_s] = M_s$ , for all  $0 \leq s \leq t \leq T$   
Interpretation. The best prediction given the information up until now is exactly the current value. This is our definition of a ‘fair game’.  
**check whether**  $up + d - dp = 0$ , **if not, not a margingale**. (probabilities of up-down with the values)  
Market is arbitrage free only if  $d < r_f < u$

In this one-period binomial model, this means we search for a measure  $\mathbb{P}^*$  identified by  $(p^*, 1 - p^*)$  such that

$$p^* \frac{S_0(1+u)}{1+r} + (1-p^*) \frac{S_0(1+d)}{1+r} = S_0$$

This is equivalent to

$$p^*(1+u) + (1-p^*)(1+d) = 1+r.$$

## Binomial model  
Solving for  $p^*$  from above:

$$p^* = \frac{r-d}{u-d} \Leftrightarrow 1-p^* = \frac{u-r}{u-d};$$

### In the two period tree

$$\begin{aligned} 100 &= 200p_{\text{u}}^* + 50p_{\text{d}}^* \Leftrightarrow p_{\text{u}}^* = \frac{1}{3}, p_{\text{d}}^* = \frac{2}{3} \\ 200 &= 300p_{\text{uu}}^* + 150p_{\text{ud}}^* \Leftrightarrow p_{\text{uu}}^* = \frac{1}{3}, p_{\text{ud}}^* = \frac{2}{3} \\ 50 &= 60p_{\text{du}}^* + 20p_{\text{dd}}^* \Leftrightarrow p_{\text{du}}^* = \frac{3}{4}, p_{\text{dd}}^* = \frac{1}{4} \end{aligned}$$

Where  $p_u + p_d = 1$ ;  $p_{uu} + p_{ud} = 1$ ;  $p_{du} + p_{dd} = 1$

### 6.3 Contingent claims (basically options)

$$\text{Pay-off CALL} \rightarrow (S_T - K)^+ = \max(0, S_T - K)$$

$$\text{Pay-off PUT} \rightarrow (S_T - K)^+ = \max(0, K - S_T)$$

*Contingent claim* with maturity  $T$  is attainable(replicable) if there exists a self-financing strategy  $\xi$  whose terminal port. value coincides with  $C_T$ , that is

$$C_T = \xi_T \cdot \bar{S}_T \quad \mathbb{P}\text{-a.s.}$$

$C_T$  is attainable if and only if its corresponding discounted claim  $H_T = \frac{C_T}{(1+r)^T}$  can be written as

$$H_T = \xi_T \cdot \bar{X}_T = V_T = V_0 + \sum_{t=1}^T \xi_t \cdot (X_t - X_{t-1})$$

for a self-financing trading strategy  $\xi$  with value process  $V$ .

The replicating strategy  $(\xi_1, \xi_2)$  must satisfy the following: 1. The self-financing property  $\xi_1 \cdot \bar{S}_1 = \xi_2 \cdot \bar{S}_1$ :

$$\begin{aligned} \xi_1^0 \cdot B_1 + \xi_1^1 \cdot S_1(\omega_u) &= \xi_2^0 \cdot B_1 + \xi_2^1 \cdot S_1(\omega_u), \\ \xi_1^0 \cdot B_1 + \xi_1^1 \cdot S_1(\omega_d) &= \xi_2^0 \cdot B_1 + \xi_2^1 \cdot S_1(\omega_d), \end{aligned}$$

2. The replicating property  $C_2 = \xi_2 \cdot \bar{S}_2$ :

$$\xi_{2,u}^0 \cdot B_2 + \xi_{2,u}^1 \cdot S_2(\omega_{uu}) = C_2(\omega_{uu}),$$

$$\begin{aligned} \xi_{2,u}^0 \cdot B_2 + \xi_{2,u}^1 \cdot S_2(\omega_{ud}) &= C_2(\omega_{ud}), \\ \xi_{2,d}^0 \cdot B_2 + \xi_{2,d}^1 \cdot S_2(\omega_{du}) &= C_2(\omega_{du}), \\ \xi_{2,d}^0 \cdot B_2 + \xi_{2,d}^1 \cdot S_2(\omega_{dd}) &= C_2(\omega_{dd}) \end{aligned}$$

## Arbitrage free prices

The goal is to price a discounted contingent claim  $H_T$  without introducing arbitrage. If  $H_T$  is attainable, the discounted initial investment needed for replicating  $H_T$  is  $\xi_1 \cdot \bar{X}_0 = V_0 = \mathbb{E}_{\mathbb{P}^*}[H_T]$ . It can be interpreted as the unique discounted arbitrage-free price for  $H$ .

**6.4 Put-call parity** At time  $T$ , one can observe that

$K + (S_T - K)^+ - (K - S_T)^+ = S_T$  This implies that  $V_T(P_1) = V_T(P_2)$ .

\*\* Put call parity\*\*

$$C_t - P_t = S_t - \frac{KB}{(1+r)^t} \Leftrightarrow KB_t + C_t - P_t = S_t$$

### Price of a european call:

$$Price(call) = \frac{1}{1+r} \left( \frac{r-d}{u-d} ((1+u)S_0 - K)^+ + \frac{u-r}{u-d} ((1+d)S_0 - K)^+ \right)$$

**6.5 Complete markets** A multiperiod arbitrage-free market is called complete if every single contingent claim is attainable

$\Rightarrow$  arbitrage-free model is complete iff there exists **just one equivalent martingale measure** (unique  $\mathbb{P}$ )

### 6.6 Replication

Formal description of payoff (left) with the replication (right). We again search for the weights  $\xi_0, \xi_1$ .

$$(S_1 - K)^+ = \xi^0 B_1 + \xi^1 S_1 = \xi^0(1+r) + \xi^1 S_1$$

$$\begin{aligned} \xi^0 &= \frac{(1+u)(S_0(1+d)-K)^+ - (1+d)(S_0(1+u)-K)^+}{(1+r)(u-d)}, \\ \xi^1 &= \frac{(S_0(1+u)-K)^+ - (S_0(1+d)-K)^+}{S_0(u-d)} \end{aligned}$$

### 6.7 T-period binomial model

$$S_t(\omega) = S_0 \cdot (1+u)^{j_t(\omega)} (1+d)^{t-j_t(\omega)},$$

where  $j_t(\omega)$  is the number of up-moves in a total of  $t$  moves, when the event  $\omega$  occurs. We can rewrite simple return as

$$\begin{aligned} R_t(\omega) &= \frac{S_t(\omega) - S_{t-1}(\omega)}{S_{t-1}(\omega)} = (1+d) \frac{1 - Y_t(\omega)}{2} + (1+u) \frac{1 + Y_t(\omega)}{2} - 1 \\ &= \frac{1}{2} (d(1 - Y_t(\omega)) + u(1 + Y_t(\omega))) = \begin{cases} u, & \text{if } Y_t(\omega) = 1 \\ d, & \text{if } Y_t(\omega) = -1 \end{cases} \end{aligned}$$

The price process of the risky asset can be written as

$$S_t = S_0 \prod_{s=1}^t (1 + R_s)$$

The discounted price process is of the form

$$X_t = \frac{S_t}{B_t} = S_0 \prod_{s=1}^t \frac{1 + R_s}{1 + r} = \frac{S_0}{(1+r)^t} \prod_{s=1}^t (1 + R_s).$$

## 7 Stochastic calculus and B-S model

- discrete time stoch. procs. (binomial trees) are approximation of the continuous time stoch. process.
- continuous time models are **limit** of discrete time models when “n” goes to infinity

**7.1 Brownian motion** Definition. A one-dimensional standard Brownian motion is a stochastic process  $(W_t)_{0 \leq t \leq T}$  with the following properties:

- $W_0 = 0$ ,
- $t \mapsto W_t(\omega)$  is a continuous function  $\mathbb{P}$ -a.s..
- $(W_t)_{0 \leq t \leq T}$  is adapted to the filtration  $(\mathcal{F}_t)_{0 \leq t \leq T}$ ,
- !!!! the increments  $W_t - W_s$  are independent and normally distributed with **variance**  $t-s$  **and zero mean** for any  $0 \leq s \leq t$ .  
D Intuitively, the change of Brownian motion on the interval  $(t, t+\Delta t)$  can be perceived as  $\Delta W_t := W_{t+\Delta t} - W_t = \epsilon \sqrt{\Delta t}$ , where  $\epsilon \sim \mathcal{N}(0, 1)$ .

#### Important properties of Brownian motion

Martingale: Given the information up to  $s < t$  the conditional expectation of  $W_t$  is  $W_s$ , that is

$$\mathbb{E}(W_t | \mathcal{F}_s) = W_s$$

Q Quadratic variation: If we divide up the time interval  $[0, t]$  into a partition  $P$  with  $n+1$  points  $t_i$ , then

$$\langle W \rangle_t = \lim_{n \rightarrow \infty} \sum_{t_i \in P} \left( W_{t_i} - W_{t_{i-1}} \right)^2 \xrightarrow{\text{a.s.}} t$$

Markov property: The conditional distribution of  $W_t$  given information up until  $s < t$  depends only on  $W_s$ . That is, for any event  $A$  and  $s < t$  it holds that

$$\mathbb{P}(W_t \in A | \mathcal{F}_s) = \mathbb{P}(W_t \in A | W_s)$$

Normality: Over finite time increments  $t_{i-1}$  to  $t_i$ ,  $W_{t_i} - W_{t_{i-1}}$  is normally distributed with mean zero and variance  $t_i - t_{i-1}$ .  
Continuity. The paths of Brownian motion are continuous, there are no discontinuities. Brownian motion is the continuous-time limit of the discrete-time random walk.  
Non-differentiability: Almost surely (with probability 1), the paths are not differentiable at any point. Therefore, it is not possible to define  $\int_0^t f(s) dW_s$  as  $\int_0^t f(s) \frac{dW_s}{ds} ds$ , i.e.

### 7.2 Stochastic integration

Stochastic process:

Definition. A process  $(H_t)_{0 \leq t \leq T}$  is a simple process if it can be written as

$$H_t(\omega) = \sum_{i=1}^p \phi_i(\omega) \mathbb{1}_{(t_{i-1}, t_j]}(t),$$

where  $0 = t_0 < t_1 < \dots < t_p = T$  and  $\phi_i$  is  $\mathcal{F}_{t_{i-1}}$ -measurable.

Stochastic integration:

Definition. A stochastic integral of a simple process  $H = (H_t)_{0 \leq t \leq T}$  is a process  $(\int_0^t H_s dW_s)_{0 \leq t \leq T}$  given by

$$\int_0^t H_s dW_s = \sum_{i=1}^p \phi_i (W_{t_i \wedge t} - W_{t_{i-1} \wedge t})$$

where  $a \wedge b = \min\{a, b\}$ .

- Note that in particular  $\int_0^t 1 dW_s = W_t$ .

- Given a  $(\mathcal{F}_t)_{0 \leq t \leq T}$ -Brownian motion  $(W_t)_{0 \leq t \leq T}$  and a  $(\mathcal{F}_t)_{0 \leq t \leq T}$ -adapted process  $(H_t)_{0 \leq t \leq T}$ , one is able to define the stochastic integral  $(\int_0^t H_s dW_s)_{0 \leq t \leq T}$  as soon as  $\int_0^T H_s^2 dW_s < \infty$ .

- the general concept extends that for simple processes.

**7.3 Ito’s lemma** Definition. Let  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$  be a filtered probability space and  $(W_t)_{0 \leq t \leq T}$  a  $\mathbb{F}$ -Brownian motion. An  $\mathbb{R}$ -valued process  $(X_t)_{0 \leq t \leq T}$  is called an Itoprocess if it can be written as

$$X_t = X_0 + \int_0^t K_s ds + \int_0^t H_s dW_s$$

where  $\geq X_0$  is  $\mathcal{F}_0$ -measurable,  
-  $(K_t)_{0 \leq t \leq T}$  and  $(H_t)_{0 \leq t \leq T}$  are  $(\mathcal{F}_t)_{0 \leq t \leq T}$ -adapted processes,  
-  $\int_0^t |K_s| ds < \infty$   $\mathbb{P}$ -a.s.,  
-  $\int_0^t H_s^2 ds < \infty$   $\mathbb{P}$ -a.s.  
Derivative “shorthand” version:  $dX_t = K_t dt + H_t dW_t$

- $dX_t$  is an increment of the process  $X_t$
- Brownian motion itself is and Ito process, where  $W_0$  is zero,  $K_s = 0$  and  $H_s = 1$

All this yields:

$$dY_t = \frac{\partial f}{\partial t}(t, X_t) dt + \frac{\partial f}{\partial x}(t, X_t) dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, X_t) d\langle X \rangle_t, \quad (3)$$

where the differentials are formally computed according to the following rules

$$dt \cdot dt = dt \cdot dW_t = dW_t \cdot dt = 0 \quad \text{and} \quad dW_t \cdot dW_t = dt$$

**Financial relevance of Ito's lemma:** Think of  $X_t$  as some underlying financial asset and of  $Y_t = f(t, X_t)$  as a new product obtained from the underlying by a possibly nonlinear transformation  $f$ . Then the Ito lemma (3) shows us how the financial product reacts to changes in the underlying. The important message of Ito's lemma is then that when using stochastic models (for  $X_t$ ), a simple linear approximation is not good enough, one must also account for the second order behavior of  $X_t$ . The trajectories of Brownian motion are of infinite variation (and finite quadratic variation) and the classical chain rule does not apply. One can view Ito's formula as a purely analytical result which provides an extension of the chain rule for  $f \circ x$  to functions  $x$  that have a nonzero quadratic variation.

## 7.4 Black-Scholes Model

### Assumptions on the assets

- $r_f$  is constant, known and same for all maturities
- log return of stock prices follows an Ito process** with constant drift ( $\mu$ ) and volatility (Normal dis.) (also implies Geometric brownian motion)
- considering only Euro options and no dividends

### Assumptions on the market

- arbitrage free, complete, perfectly liquid and divisible
- frictionless (No TC)
- delta hedging is performed continuously

### Formula of B-S derived by Ito

$$S_t = S_0 \cdot \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right)$$

If we want to achieve **zero randomness in self financing portfolio**, which looks like

$$\text{this: } d\Pi_t = \frac{\partial V}{\partial t}(t, S_t) dt + \frac{\partial V}{\partial S}(t, S_t) dS_t + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2}(t, S_t) dt - \theta_t dS_t$$

Then we have to choose the proper weight theta of the short leg so the gain offset each other. This is achieved when:

$$\theta_t = \frac{\partial V}{\partial S}(t, S_t) := \Delta_V$$

- rebalanced continuously: dynamic delta hedging

After setting up the initial portfolio, the first der of the portfolio value is:

$$d\Pi_t = \left( \frac{\partial V}{\partial t}(t, S_t) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2}(t, S_t) \right) dt.$$

### Black scholes Partial differential equaiton (PDE)

We should know the proof. Consult L06,s.37-41.

B-S PDE =

$$\frac{\partial V}{\partial t}(t, S_t) + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2}(t, S_t) + rS_t \frac{\partial V}{\partial S}(t, S_t) - rV(t, S_t) = 0$$

- B-S does not depend on  $\mu$  !!

## 7.5 Black-scholes by Martingale

This yields the Black-Scholes formula for call options:

$$C(t, S_t) = S_t \Phi(d_1) - K e^{-r(T-t)} \Phi(d_2)$$

where

$$d_1 = \frac{\ln\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}$$

and

$$d_2 = d_1 - \sigma\sqrt{T-t}$$

$\Phi(\cdot)$  denotes the cdf of a  $\mathcal{N}(0, 1)$  distributed random variable

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{s^2}{2}} ds$$

Alternatively, we can derive it using the put-call parity relation

$$C(t, S_t) - P(t, S_t) = S_t - K \cdot e^{-r(T-t)}.$$

The put option price is

$$P(t, S_t) = K \cdot e^{-r(T-t)} \cdot \Phi(-d_2) - S_t \cdot \Phi(-d_1).$$

## 8 Risk measures

**Risk** is "any event or action that may adversely affect an organization's ability to achieve its objectives and execute its strategies"

### Approaches to Risk measurement:

- notional-amount approach: the risk of a portfolio is defined as the (weighted) sum of the notional values of the individual securities
- factor sensitivity measures: give the change in portfolio value for a given predetermined change in one of the underlying risk factors
- scenario-based risk measures: one considers a number of future scenarios and measures the maximum loss of the portfolio under these scenarios
- risk measures based on loss distributions: stat. quantities describing characteristics of the loss distribution of the portfolio

In our (CARM) interpretation, risk measures how much capital to inject to become adequately capitalized (under some regulations)

### Three pillar concept under Basel 2

- minimal capital charge: requirements for the calculation of the regulatory capital to ensure that a bank holds sufficient capital for its credit risk in the banking book its market risk in the trading book and operational risk
- supervisory review process: local regulators review the check and balances put in place for capital adequacy assessments, ensure that banks have adequate reg. capital and perform stress tests of a bank's capital adequacy
- market discipline: banks are required to make their risk management processes more transparent

### Reasons for Corp Risk management (turning M&M upside down)

- RM can reduce taxes
- RM is beneficial coz of better access to cap. markets
- RM increases value in presence of bankruptcy costs (lower expected value coz of lower P(default))
- RM reduces impact of costly external financing

### 8.1 Acceptance set

Definition. A set  $\mathcal{A}$  of random variables is called an acceptance set if

$\triangleright \mathcal{A}$  is non-trivial, i.e.  $\mathcal{A} \neq \emptyset$  and  $\mathcal{A} \subsetneq \mathcal{X}$ ,  
 $\triangleright \mathcal{A}$  is monotone, i.e.  $X \in \mathcal{A}, Y \geq X \implies Y \in \mathcal{A}$  (if  $X$  is acceptable and  $Y = X + 100$ ,  $X$  is also acceptable)

By  $Y \geq X$  we mean either  $Y(\omega) \geq X(\omega)$  for all  $\omega \in \Omega$ , or  $\mathbb{P}$ -almost surely. The acceptance set is specified by the regulator.

An acceptance set is called: (L08, 17 for pictures)

- cone* if  $X_T \in \mathcal{A} \implies \forall \lambda > 0: \lambda X_T \in \mathcal{A}$
- convex* if  $X_T, Y_T \in \mathcal{A} \implies \forall \lambda \in [0, 1]: \lambda X_T + (1 - \lambda)Y_T \in \mathcal{A}$
- closed* if  $\mathcal{A} = \overline{\mathcal{A}}$ ,

### 8.2 VaR

$$\text{VaR}_\alpha(X) = -\mu - \sigma \phi^{-1}(\alpha)$$

- Finding var: order the outcomes (returns/cap positions) in a decreasing order. Start from the first and add probabilities until they are **strictly greater than  $\alpha$** . Once they are, take the corresponding  $X_n \cdot -1$  and that is VaR at level alpha.

### Remarks

- VaR does not account for severity of losses with prob  $1 - \alpha$
- VaR does not account for diversification
- VaR is not convex, hence not coherent - might penalize diversification
- Simon Johnson (MIT): "VaR misses everything that matters when it matters"

### 8.3 Expected shortfall ES

$$\mathcal{A}_{\text{ES}\alpha} := \left\{ X \in \mathcal{X} \mid \frac{1}{\alpha} \int_0^\alpha \text{VaR}_\beta(X) d\beta \leq 0 \right\}$$

$$\text{ES}_\alpha(X) = -\mu + \frac{\sigma}{\alpha} \Phi'(\Phi^{-1}(\alpha))$$

An acceptance set  $\mathcal{A}$  is said to be coherent if it is

- convex:  $X, Y \in \mathcal{A}, 0 < \lambda < 1 \implies \lambda X + (1 - \lambda)Y \in \mathcal{A}$ ,

- conic:  $X \in \mathcal{A}, \lambda \geq 0 \implies \lambda X \in \mathcal{A}$

Financial interpretation:

- Convexity means that portfolios of acceptable positions are still acceptable, i.e. acceptability is preserved by diversification.

- Conicity means that acceptability is independent of the position size.

## 9 Mean-variance Portfolio/Optimization MVP

- we are risk averse but like money. Described by utility functions. Second deriv is  $\leq 0$ . Thus, function must be increasing and concave (slide 5)

## 9.1 Diversification principle

$$X = \begin{cases} 2m & \text{coin shows heads} \\ 0 & \text{coin shows tails.} \end{cases} \quad Y = \begin{cases} 0 & \text{coin shows heads} \\ 2m & \text{coin shows tails.} \end{cases}$$

Play both at once. Invest  $\frac{m}{2}$  in either game, total of  $m$  invested. The new payoff is given by

$$\frac{1}{2}X + \frac{1}{2}Y = \begin{cases} m & \text{coin shows heads} \\ m & \text{coin shows tails.} \end{cases}$$

So in either case, we get our money back! The payoff is constant. Hence,

$$\begin{aligned} \text{Expectation:} \quad \mathbb{E}\left[\frac{1}{2}(X + Y)\right] &= \frac{1}{2} \cdot m + \frac{1}{2} \cdot m = m \\ \text{Variance:} \quad \mathbb{V}\left[\frac{1}{2}(X + Y)\right] &= \frac{1}{2} \cdot m^2 + \frac{1}{2} \cdot m^2 - m^2 = 0 \end{aligned}$$

We achieved diversification because of perfect negative correlation, which is a necessary, but not a sufficient condition.

$$\begin{aligned} \rho_{X,Y} &:= \text{corr}(X, Y) := \frac{\text{Cov}(X, Y)}{\sqrt{\mathbb{V}[X]}\sqrt{\mathbb{V}[Y]}} := \frac{\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]}{m \cdot m} \\ &= \frac{\mathbb{E}[XY - mX - mY + m^2]}{m^2} = \frac{\mathbb{E}[XY] - m^2 + m^2 + m^2}{m^2} \\ &= \frac{0 - m^2}{m^2} = -1 \end{aligned}$$

## 9.2 Always to remember assumptions

- preference to increase consumption

- risk-aversion

- therefore, utility functions are concave

- an investor maximises the portfolio's expected return for a given level of variance or vice versa, minimises the variance for a given level of expected return.**

## 9.3 Var-Covar matrix

- Finally, we need one more ingredient. The covariance between two returns  $\ell, m \in \{1, \dots, M\}$  is defined as  $\text{Cov}(R_{T,\ell}, R_{T,m}) = \sum_{n=1}^N p_n (R_{T,\ell}(\omega_n) - \mathbb{E}[R_{T,\ell}]) (R_{T,m}(\omega_n) - \mathbb{E}[R_{T,m}])$ . We assemble the individual covariances in the covariance matrix  $\text{Cov}(\mathbf{R}_T) \in \mathbb{R}^{M \times M}$  of the random vector,  $\text{Cov}(\mathbf{R}_T) = \begin{bmatrix} \text{Cov}(R_{T,1}, R_{T,1}) & \text{Cov}(R_{T,1}, R_{T,2}) & \cdots & \text{Cov}(R_{T,1}, R_{T,M}) \\ \text{Cov}(R_{T,2}, R_{T,1}) & \text{Cov}(R_{T,2}, R_{T,2}) & \cdots & \text{Cov}(R_{T,2}, R_{T,M}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(R_{T,M}, R_{T,1}) & \text{Cov}(R_{T,M}, R_{T,2}) & \cdots & \text{Cov}(R_{T,M}, R_{T,M}) \end{bmatrix}$ . Note that  $\text{Cov}(R_{T,m}, R_{T,m}) = \mathbb{V}[R_{T,m}]$ , and thus, the diagonal of  $\text{Cov}(\mathbf{R}_T)$  is the variance vector,  $\text{diag}(\text{Cov}(\mathbf{R}_T)) = \mathbb{V}[\mathbf{R}_T]$ .

**9.4 Implementation portfolio** is a *linear combination* of all individual assets. Coefficient  $\alpha$  denote the weights. We investing all our wealth, i.e. sum of alphas = 1  
Simple return of the portfolio:

$$\begin{aligned} R_T^\alpha &= \frac{P_T^\alpha}{P_0^\alpha} - 1 = \frac{\sum_{m=1}^M \frac{\alpha_m W}{P_{0,m}} P_{T,m}}{W} - \sum_{m=1}^M \alpha_m \\ &= \sum_{m=1}^M \alpha_m \underbrace{\left( \frac{P_{T,m}}{P_{0,m}} - 1 \right)}_{=R_{T,m}} = \sum_{m=1}^M \alpha_m R_{T,m}. \end{aligned}$$

$$\mathbb{E}[R_T^\alpha] = \sum_{n=1}^N p_n R_T^\alpha(\omega_n)$$

**Portfolio's return as linear combination of individual expectations.**  
Showing that expectation of the portfolio's returns  
Based on the result from the previous exercise, we can write  $R_T^\alpha$  as the linear combination of the individual returns,

$$\begin{aligned}\mathbb{E}\left[R_T^\alpha\right] &= \sum_{n=1}^N p_n R_T^\alpha\left(\omega_n\right)=\sum_{n=1}^N p_n \sum_{m=1}^M \alpha_m R_{T, m}\left(\omega_n\right) \\ &= \sum_{m=1}^M \alpha_m \underbrace{\sum_{n=1}^N p_n R_{T, m}\left(\omega_n\right)}_{=\mathbb{E}\left[R_{T, m}\right]}=\sum_{m=1}^M \alpha_m \mathbb{E}\left[R_{T, m}\right]=\boldsymbol{\alpha}^T \mathbb{E}\left[\mathbf{R}_T\right].\end{aligned}$$

§  
**Variance of the portfolio's returns (proof s.59)**

$$\mathbb{V}\left[R_T^\alpha\right]=\sum_{\ell=1}^M \sum_{m=1}^M \alpha_\ell \alpha_m \operatorname{Cov}\left(R_{T, \ell}, R_{T, m}\right)=\boldsymbol{\alpha}^\top \operatorname{Cov}\left(\mathbf{R}_T\right) \boldsymbol{\alpha}$$

**Constrained optimasation (Markowitz' port. optimization problem / MVO)** M-V is a **constrained optimisation problem**:

- minimise or maximise an objective function over all possible portfolio combinations, hence, over all possible weight vectors  $\alpha$
- The optimisations are subject to a set of constraints

$$\mathcal{C}$$

**First way: minimize risk for set return MV-Min** We only have a budget (wealth) and an objective constraint (fixing return). Arbitrary short-selling is allowed.

$$\mathcal{C}_E=\left\{\mathbf{1}^\top \boldsymbol{\alpha}=1, \quad \mathbb{E}\left[R_T^\alpha\right] \geq E^* \text { for some fixed } E^* \in \mathbb{R}\right\}; \text { where } E^* \text { is our required return}$$

Choosing too large required return can lead to empty set. In words, the first constraint (budget) requires alphas to sum up to 1. The second say expected return is no less than the target.

**Second way: maximize return for set risk MV-Max**

$$\mathcal{C}_V=\left\{\mathbf{1}^\top \boldsymbol{\alpha}=1, \quad \mathbb{V}\left[R_t^\alpha\right] \leq V^* \text { for some fixed } V^*>0\right\} .$$

**Example**

Stock	Expected Return	Standard Deviation
A	$\mu_A=10\%$	$\sigma_A=5\%$
B	$\mu_B=15\%$	$\sigma_B=10\%$

Correlation  $\rho=-1$

$$\mu=\left(\begin{array}{c} \mu_A \\ \mu_B \end{array}\right)=\left(\begin{array}{c} 0.1 \\ 0.15 \end{array}\right) \text { and } \Sigma=\left(\begin{array}{cc} \sigma_A^2 & \sigma_A \sigma_B \rho \\ \sigma_A \sigma_B \rho & \sigma_B^2 \end{array}\right)=\left(\begin{array}{cc} 0.0025 & -0.005 \\ -0.005 & 0.01 \end{array}\right)$$

Setting the weights of the portfolio  $\omega=\left(\omega_A, \omega_B\right)^\top$ , the variance of the portfolio is given by

$$\begin{aligned}\sigma_p^2 &= \omega^\top \Sigma \omega = \left(\begin{array}{cc} \omega_A & \omega_B \end{array}\right) \left(\begin{array}{cc} \sigma_A^2 & \sigma_A \sigma_B \rho \\ \sigma_A \sigma_B \rho & \sigma_B^2 \end{array}\right) \left(\begin{array}{c} \omega_A \\ \omega_B \end{array}\right) \\ &= \omega_A^2 \sigma_A^2 + 2 \omega_A \omega_B \sigma_A \sigma_B \rho + \omega_B^2 \sigma_B^2 \\ &= \omega_A^2 \cdot 0.0025 - \omega_A \omega_B \cdot 0.01 + \omega_B^2 \cdot 0.01 \\ &= \left(0.05 \omega_A - 0.1 \omega_B\right)^2\end{aligned}$$