Friedrich-Alexander-Universität Erlangen-Nürnberg



Decision Theory

Lecture 10

Michael Hartisch

Friedrich-Alexander Universität Erlangen-Nürnberg, Department Data Science June 24, 2024



Recap: What did we do?

- Decision under risk
- μ -principle
- μ - σ -principle
- Functions based on other measures

FAU M. Hartisch Decision Theory June 24, 2024 2/61



Decisions Under Risks

• Know probability p_j for each scenario s_j



μ -Prinziple

Decide only based on the expected value

$$\mu_i = \sum_{j \in [n]} p_j e_{ij}$$

- Pro:
 - Law of large numbers
 - If decision is repeated, this is the best we can do
- Cons:
 - Decision once, law of large numbers does not hold
 - Never play lottery or take insurance
 - Petersburg paradox



Petersburg Paradox

• How good is this game?

Expected profit:

$$\mu = 2 \cdot \frac{1}{2} + 4\frac{1}{4} + 8 \cdot \frac{1}{8} + \dots = \infty$$

Clearly not how we decide



Certainty Equivalent

Value CE such that

$$\Phi(CE) = \Phi(a_i)$$

- For how much are you willing to sell your ticket?
- In the Petersburg paradox: infinite CE



μ - σ -Principle

• Include both expected value μ and standard deviation σ (or variance σ^2)

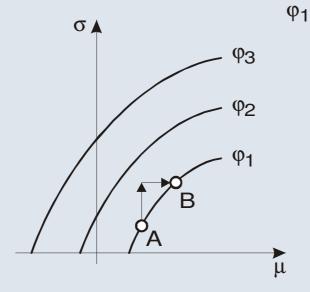
$$\sigma = \sqrt{\sum_{j=1}^n p_j (x_j - \mu)^2}$$

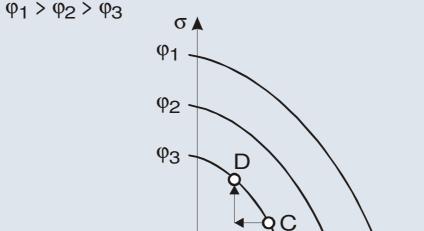
$$\sigma^2 = \sum_{j=1}^n p_j (x_j - \mu)^2$$

- $\partial \Phi(\mu, \sigma)/\partial \sigma > 0$: risk-seeking
- $\partial \Phi(\mu, \sigma)/\partial \sigma = 0$: risk-neutral
- $\partial \Phi(\mu, \sigma)/\partial \sigma < 0$: risk-averse



Level Sets





- Draw line through points which give same value
- Map of a landscape: points on the same height
- On the left: risk-averse
- On the right: risk-seeking



μ - σ Preference Functions

$$\Phi(\mu, \sigma) = \mu - \sigma^2$$

- Looks reasonable: compromise between expected profit and risk
- But can choose dominated solutions!



Functions Based on other Measures

- μ and σ are used to represent the distribution of results
- Many more measures exist in statistics
- Depending on what is important for the decision at hand, one might rather want to use one of these
- Examples:
 - Worst case
 - Quantiles
 - Loss probability
 - Loss expectation
 - Range of variation



Today

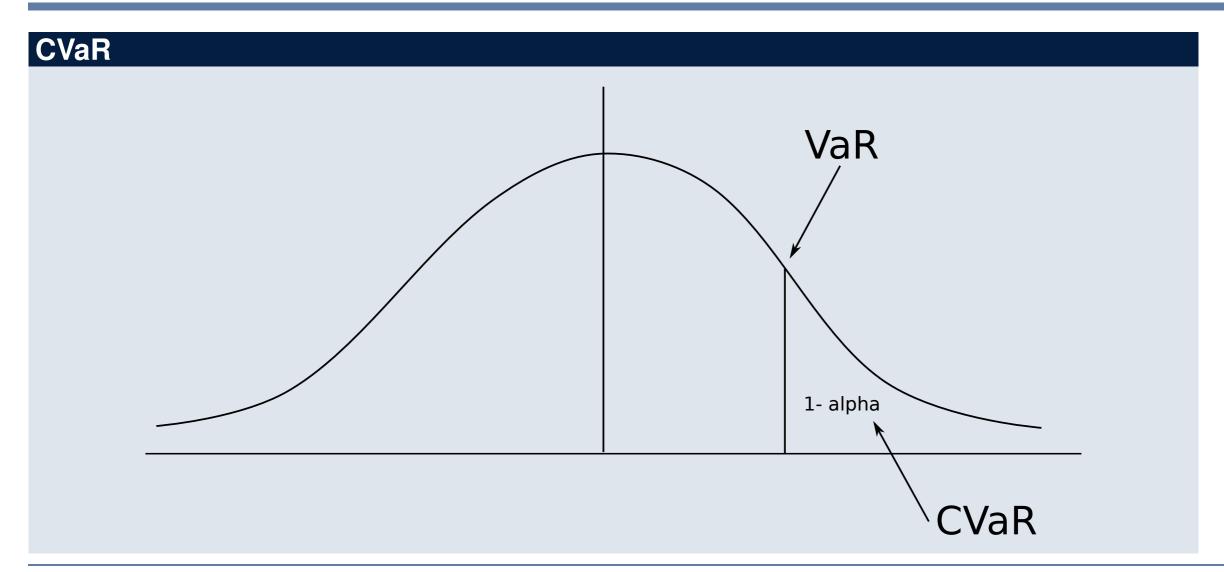
- CVaR
- Application example: portfolio planning
- Bernoulli principle



Conditional Value at Risk

- Extensively studied (also for portfolio): conditional value at risk (CVaR)
- also referred to as "expected shortfall"
- estimates risk of an investment; focus on less profitable outcomes (conservative measure)
- Value at risk VaR_{α} (Minimization):
 - \circ What is the value x such that the probability of being less than or equal to x is α ? (Quantile)
- Conditional value at risk $CVaR_{\alpha}$ (Minimization):
 - \circ What is the expected value of all outcomes greater than VaR_{α} ?







- What is the expected value of the 1 $-\alpha$ worst outcomes?
- Minimize:

- CVaR for α = 0:
 - Expected value of all outcomes



- What is the expected value of the 1 $-\alpha$ worst outcomes?
- Minimize:

- CVaR for α = 0:
 - Expected value of all outcomes
- CVaR for α = 0.2:
 - 30% outcome 8
 - 40% outcome 6
 - 10% outcome 4
 - CVaR is 6.5



- What is the expected value of the 1 $-\alpha$ worst outcomes?
- Minimize:

- CVaR for α = 0:
 - Expected value of all outcomes
- CVaR for α = 0.5:
 - 30% outcome 8
 - o 20% outcome 6
 - o CVaR is 7.2



- What is the expected value of the 1 $-\alpha$ worst outcomes?
- Minimize:

- CVaR for α = 0:
 - Expected value of all outcomes
- CVaR for α = 0.8:
 - 20% outcome 8
 - CVaR is 8
 - \circ For α close to 1, CVaR is equal to worst case



CVaR

• What happens if everything is equally probable?

- CVaR for $\alpha = 1/4$ is the average of the 3 worst outcomes
- CVaR for $\alpha = 2/4$ is the average of the 2 worst outcomes
- CVaR for $\alpha = 3/4$ is the average of the 1 worst outcome



Connection to OWA

• Reminder: OWA with weights (1/2, 1/2, 0, 0) gives

$$\frac{1}{2}$$
(largest outcome) + $\frac{1}{2}$ (second-largest outcome)

- Therefore: for *n* scenarios (all equal probability), the following is equivalent
 - \circ OWA with $w = (1/k, 1/k, \dots, 1/k, 0, 0, \dots, 0)$ (the first k terms are 1/k)
 - CVaR with $\alpha = (n k)/n$



$$CVaR_{0.1} =$$

$$CVaR_{0.5} =$$

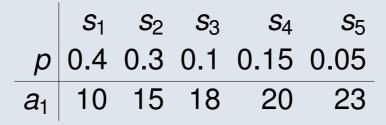
$$CVaR_{0.9} =$$

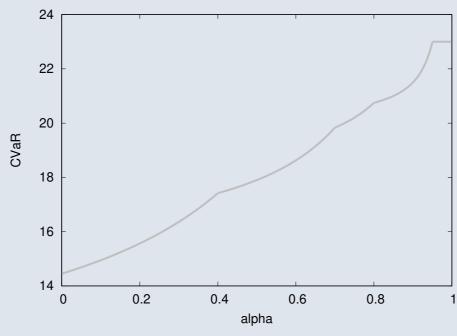


$$CVaR_{0.1} = (0.05 \cdot 23 + 0.15 \cdot 20 + 0.1 \cdot 18 + 0.3 \cdot 15 + 0.3 \cdot 10)/0.9$$

= 14.9
 $CVaR_{0.5} = (0.05 \cdot 23 + 0.15 \cdot 20 + 0.1 \cdot 18 + 0.2 \cdot 15)/0.5$
= 17.9
 $CVaR_{0.9} = (0.05 \cdot 23 + 0.05 \cdot 20)/0.1$
= 21.5









Application of $\mu - \sigma$: Portfolio Theory

- Aim to find the optimal composition of our stock portfolio
- Portfolio return e_i , j = 1, ..., n possible environmental states

$$\mu = \sum_{j \in [n]} e_{j} p_{j}$$

$$\sigma^{2} = \sum_{j \in [n]} (e_{j} - \mu)^{2} \cdot p_{j}$$

$$= \sum_{j \in [n]} e_{j}^{2} p_{j} - 2\mu \sum_{j \in [n]} e_{j} p_{j} + \mu^{2} \sum_{j \in [n]} p_{j}$$

$$= \sum_{j \in [n]} e_{j}^{2} p_{j} - \mu^{2}$$



Portfolio Example

- Let's assume there are two stocks
- Return in scenario j is e_{1j} , e_{2j}
- We hold shares x_1 and x_2
- $e_j = e_{1j}X_1 + e_{2j}X_2$
- $\mu = \sum_{j \in [n]} p_j e_j = \mu_1 x_1 + \mu_2 x_2$



Portfolio Example

- Investor wants to invest 100,000 EUR for exactly one year
- Achieve maximum return
- Applies μ - σ principle
- Decides only between stock 1 and 2
- Beginning of year 0
- Determines the annual returns of the last 5 years



Portfolio Example

Year

 -5
 -4
 -3
 -2
 -1

 Stock

$$s_1$$
 s_2
 s_3
 s_4
 s_5

 1
 15
 5
 -5
 25
 10

 2
 -10
 50
 10
 30
 20

- The investor considers all five development possibilities equally likely, $p_i = 0.2$
- μ_1 = 10, σ_1 = 10
- μ_2 = 20, σ_2 = 20



Portfolio Example

- The investor perceives the risk of a_2 as too high
- He considers two alternatives:
 - a₁: Invest entirely in Stock 1
 - a₂: Invest 90,000 EUR in stock 1 and 10,000 EUR in stock 2



Portfolio Example

- a_1 evidently has $\mu_{a_1} = 10$, $\sigma_{a_1} = 10$
- For *a*₂:

- μ_{a_2} = 11 and σ_{a_2} = 9.2
- The mixture is better in terms of expected value and standard deviation



Portfolio Example

- The expected value $e_i = e_{1j}x_1 + e_{2j}x_2$ is linear
- Therefore, the new expected value is a combination of the old ones,

$$\mu_1 = 10, \ \mu_2 = 20, \ \mu_{a_2} = 0.9 \cdot 10 + 0.1 \cdot 20 = 11$$

This does not apply to standard deviation:

$$\sigma_1 = 10, \ \sigma_2 = 20, \ \sigma_{a_2} = 9.2$$

We use a riskier stock to reduce risk!



Correlation

- Why is this so?
- Need correlation ρ
- Determines how two random variables respond similarly to changes in the environment
 - $\circ \rho$ = 1: Same
 - \circ ρ = -1: Opposite
 - $\circ \rho = 0$: Independent
- Formula:

$$\rho_{12} = \frac{cov_{12}}{\sigma_1 \cdot \sigma_2}$$

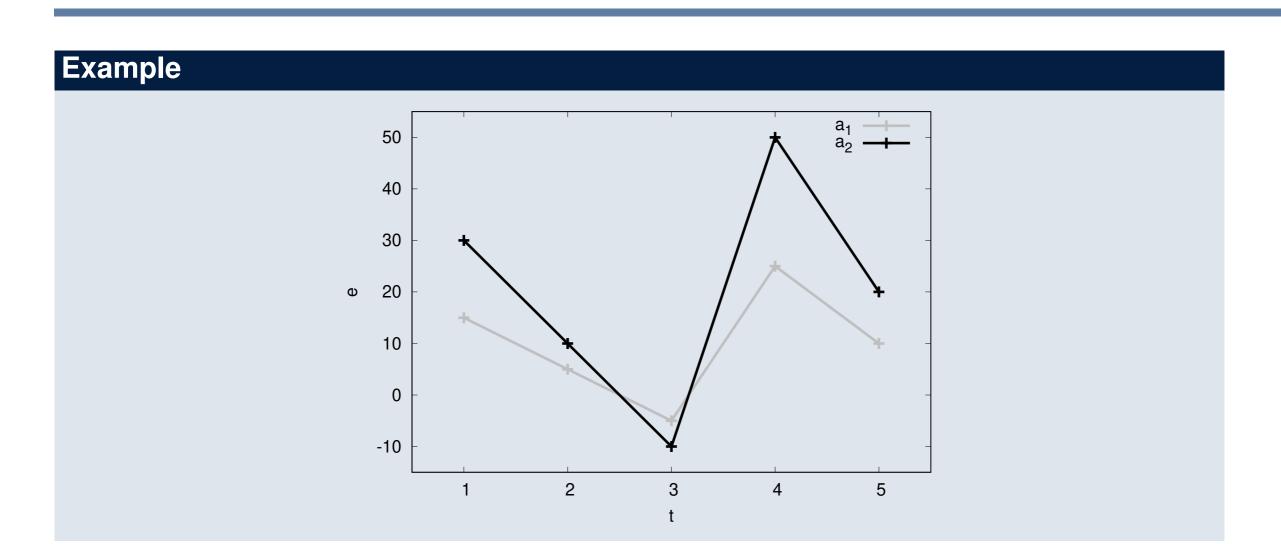
with cov_{12} as the covariance between variables 1 and 2:

$$cov_{12} = \sum_{j \in [n]} p_j(e_{1j} - \mu_1) \cdot (e_{2j} - \mu_2)$$



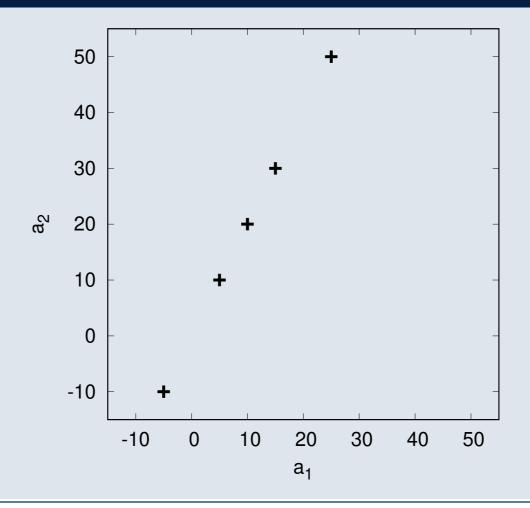
- μ_1 = 10, σ_1 = 10
- μ_2 = 20, σ_2 = 20
- $cov_{12} = \frac{1}{5}(50 + 50 + 450 + 450 + 0) = 200$
- $\rho_{12} = cov_{12}/\sigma_1\sigma_2 = 200/(10 \cdot 20) = 1$
- Stock 1 and 2 are 100% positively correlated







Example



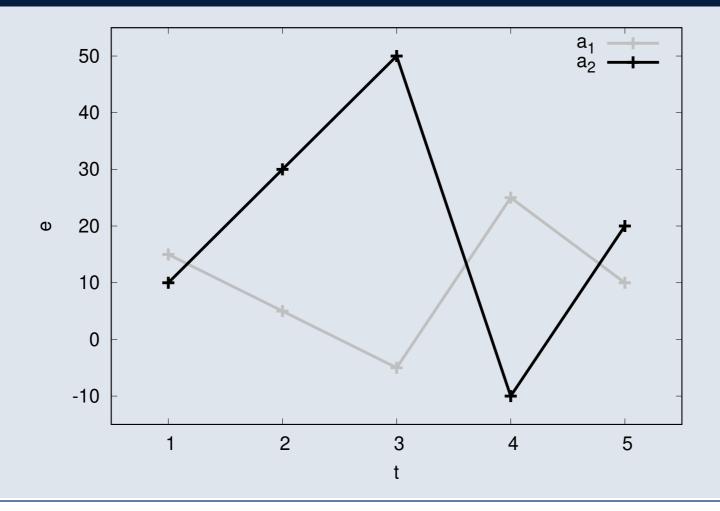
FAU M. Hartisch Decision Theory June 24, 2024 29/61



- μ_1 = 10, σ_1 = 10
- μ_2 = 20, σ_2 = 20
- $cov_{12} = -200$
- $\rho_{12} = cov_{12}/\sigma_1\sigma_2 = -200/(10 \cdot 20) = -1$
- Stock 1 and 2 are 100% negatively correlated

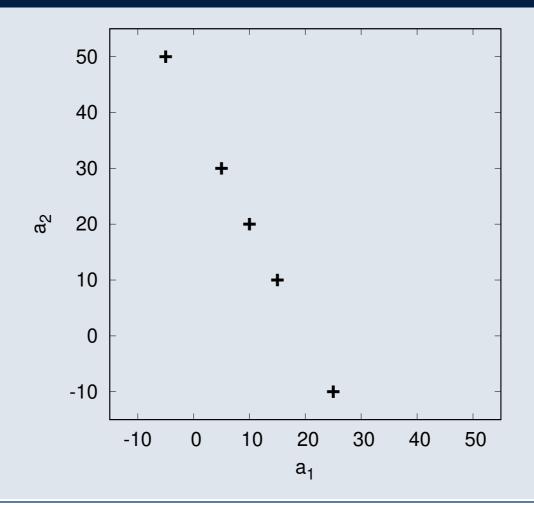


Example





Example



FAU M. Hartisch Decision Theory June 24, 2024 32/61



Example

•
$$\mu_1$$
 = 10, σ_1 = 10

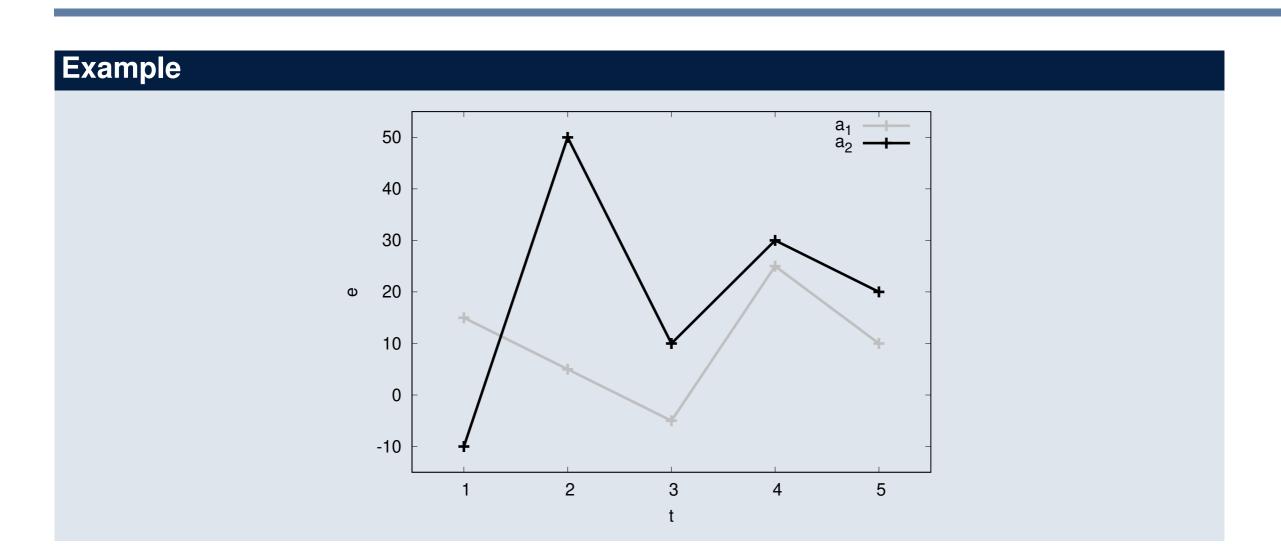
•
$$\mu_2$$
 = 20, σ_2 = 20

•
$$cov_{12} = 0$$

•
$$\rho_{12} = cov_{12}/\sigma_1\sigma_2 = 0/(10 \cdot 20) = 0$$

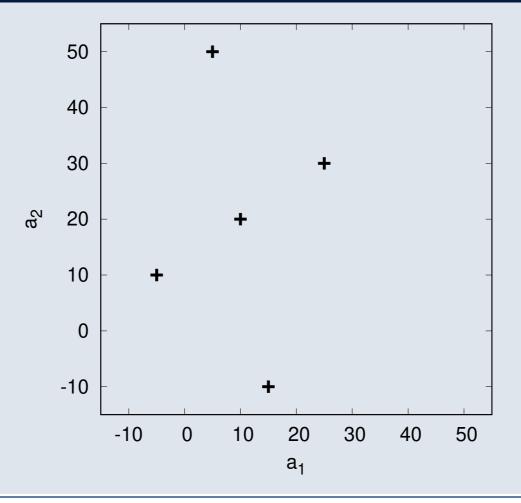
Stocks 1 and 2 are not correlated







Example





Correlation

• Depending on the proportions x_1 and x_2 , the variance of the portfolio is given by:

$$\sigma^2 = x_1^2 \sigma_1^2 + x_2^2 \sigma_2^2 + 2x_1 x_2 \sigma_1 \sigma_2 \rho_{12}$$

• In special cases:

$$\circ \rho = 1 \rightarrow \sigma^2 = (x_1\sigma_1 + x_2\sigma_2)^2$$

 \triangleright Standard deviation is linear in x_1 , x_2 (similar to expected value)

$$\circ \rho = -1 \rightarrow \sigma^2 = (x_1\sigma_1 - x_2\sigma_2)^2$$

- ► We can reduce variance by combining appropriately
- Also known as "Hedging"

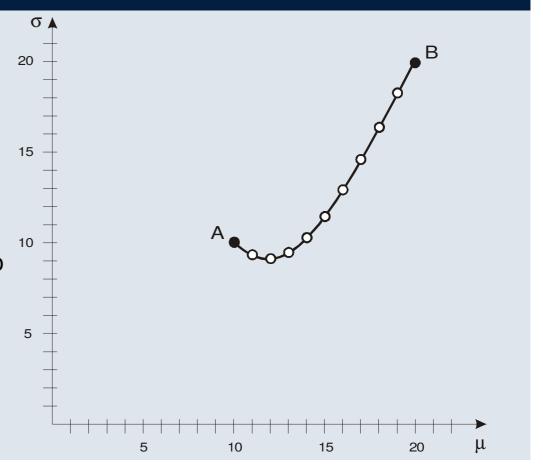
$$\rho = 0 \rightarrow \sigma^2 = x_1^2 \sigma_1^2 + x_2^2 \sigma_2^2$$

► Reduction in variance is also possible here



Correlation

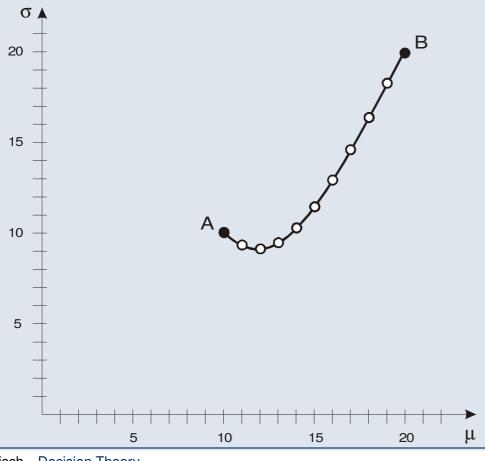
- We only need to set x_1 , and x_2 is then defined as $1 x_1$
- Consider all combinations for $x_1 = 1$ (A) to $x_1 = 0$ (B)
- \bullet Plot μ and σ in a diagram for each combination (portfolio line)



FAU M. Hartisch Decision Theory



Correlation



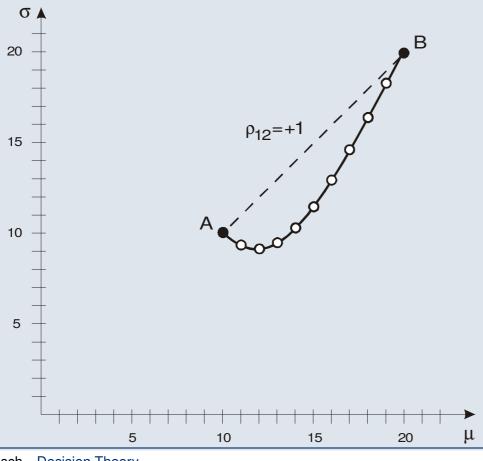
- μ - σ diagram:
- Suppose we have $\rho = 1$

$$\sigma = X_1 \sigma_1 + X_2 \sigma_2$$
$$= \sigma_2 + (\sigma_1 - \sigma_2) X_1$$

- We can only increase expected value if we increase standard deviation equally
- Linear relationship



Correlation



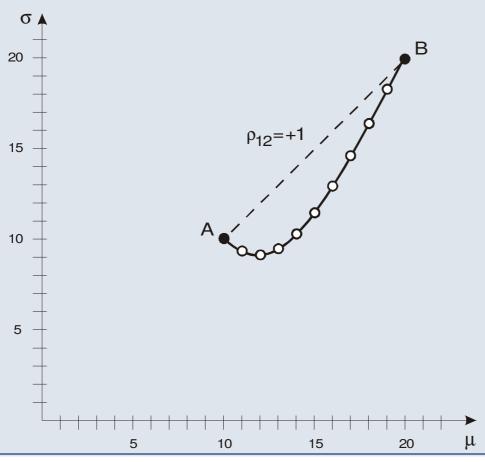
- μ - σ diagram:
- Suppose we have $\rho = 1$

$$\sigma = X_1 \sigma_1 + X_2 \sigma_2$$
$$= \sigma_2 + (\sigma_1 - \sigma_2) X_1$$

- We can only increase expected value if we increase standard deviation equally
- Linear relationship



Correlation



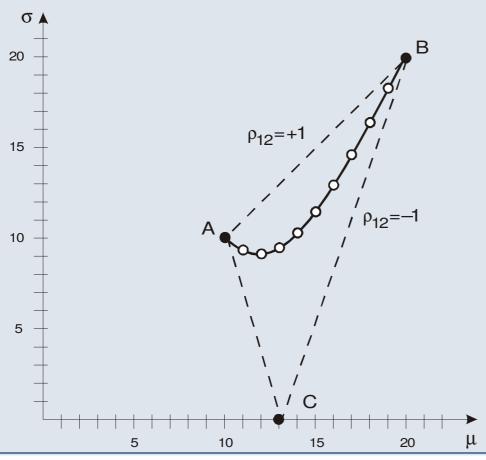
- μ - σ diagram:
- Suppose we have $\rho = -1$

$$\sigma = |x_1\sigma_1 - x_2\sigma_2|$$
$$= |(\sigma_1 + \sigma_2)x_1 - \sigma_2|$$

• For $x_1 = \sigma_2/(\sigma_1 + \sigma_2)$, we get $\sigma = 0$



Correlation



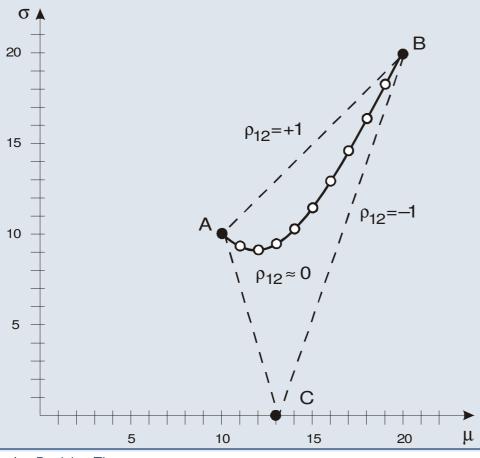
- μ - σ diagram:
- Suppose we have $\rho = -1$

$$\sigma = |x_1\sigma_1 - x_2\sigma_2|$$
$$= |(\sigma_1 + \sigma_2)x_1 - \sigma_2|$$

• For $x_1 = \sigma_2/(\sigma_1 + \sigma_2)$, we get $\sigma = 0$



Correlation



- μ - σ diagram:
- Suppose ρ is between -1 and +1
- Get a curve in triangle ABC
- For ρ close to -1: convex, close to C
- For ρ close to +1: steep, close to line AB



Example

- We continue to have Stock 1 with $\mu_1 = 10$ and $\sigma_1 = 10$ as an option
- Also, there are government bonds with a one-year maturity
- Guaranteed to have no variance
- μ_0 = 6 and σ_0 = 0
- Proportions x_0 and x_1
- Expected value?

$$\mu = 6x_0 + 10x_1$$

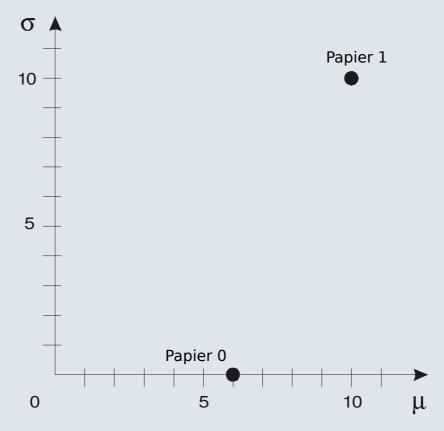
• Variance?

$$\sigma^2 = \sigma_0^2 x_0^2 + \sigma_1^2 x_1^2 + 2x_0 x_1 \sigma_0 \sigma_1 \rho_{12} = \sigma_1^2 x_1^2 = 100 x_1^2$$



Portfolio Line

• What does the portfolio line look like?



FAU M. Hartisch Decision Theory June 24, 2024 40/61



Portfolio Line

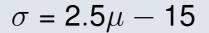
- $x_0 + x_1 = 1$
- $\mu = \mu_0 x_0 + \mu_1 x_1 = \mu_0 (1 x_1) + \mu_1 x_1 = (\mu_1 \mu_0) x_1 + \mu_0$
- $X_1 = \frac{\mu \mu_0}{\mu_1 \mu_0}$
- $\sigma = \sigma_1 X_1 = \sigma_1 \cdot \frac{\mu \mu_0}{\mu_1 \mu_0} = \frac{\sigma_1}{\mu_1 \mu_0} \mu \frac{\mu_0 \sigma_1}{\mu_1 \mu_0}$
- With μ_0 = 6, μ_1 = 10, and σ_1 = 10, we have:

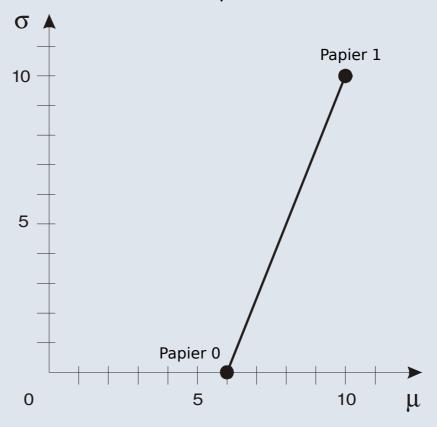
$$\sigma = 2.5\mu - 15$$

FAU M. Hartisch Decision Theory June 24, 2024 41/61



Visualization







Portfolio Optimization

- Which portfolio is the best?
- Must know the preference function $\Phi(\mu, \sigma)$
- Substitution: function in (x_1, x_2) , then only x_1
- Take the derivative with respect to x_1 and set it equal to zero

FAU M. Hartisch Decision Theory

June 24, 2024 43/61



Portfolio Optimization

- Which portfolio is the best?
- Must know the preference function $\Phi(\mu, \sigma)$
- Substitution: function in (x_1, x_2) , then only x_1
- Take the derivative with respect to x_1 and set it equal to zero

In the Example

$$\Phi(\mu, \sigma) = \mu - \alpha \sigma^{2} = (\mu_{1} - \mu_{0})x_{1} + \mu_{0} - \alpha \sigma_{1}^{2}x_{1}^{2}$$

$$\Phi'(x_{1}) = \mu_{1} - \mu_{0} - \alpha \sigma_{1}^{2}x_{1}$$

$$\Phi'(x_{1}) = 0 \Leftrightarrow x_{1} = \frac{\mu_{1} - \mu_{0}}{\alpha \sigma_{1}^{2}}$$



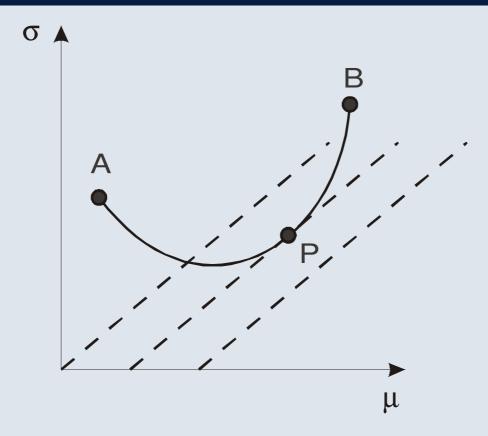
Portfolio Optimization

- Can also use a portfolio line
- $\Phi(\mu, \sigma) = \mu \alpha \sigma$, $\alpha > 0$: linear level curves
- $\Phi(\mu, \sigma) = \mu \alpha \sigma^2$, $\alpha > 0$: concave level curves
- Find the intersection point of the highest level curve with the portfolio line

FAU M. Hartisch Decision Theory



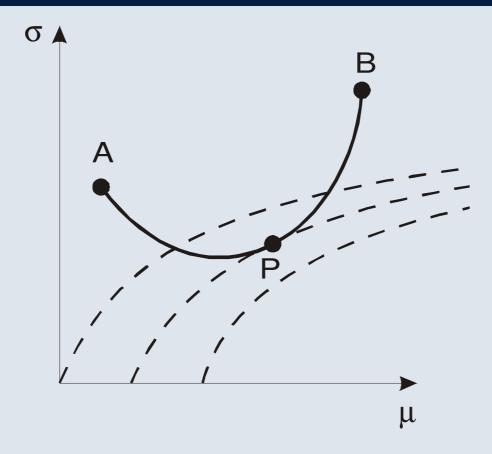
Case $\Phi(\mu, \sigma) = \mu - \alpha \sigma$ (Linear Level Curves)



The best portfolio is at point P



Case $\Phi(\mu, \sigma) = \mu - \alpha \sigma^2$ (Concave Level Curves)



The best portfolio is at point P

FAU M. Hartisch Decision Theory



Relationship

- Smaller α :
 - We are more risk-friendly
 - Level curves are steeper
 - Intersection point is further to the right
- Larger α :
 - We are more risk-averse
 - Level curves are flatter
 - Intersection point is further to the left
- With more risk, we can achieve portfolios with a better expected value

FAU M. Hartisch Decision Theory



Bernoulli Principle – What is it about?

(Daniel) Bernoulli, 1738:

- Evaluation of gambling games
- Not the expected value of the win
- "Moral expectation" = expected value of utility
- Utility not linear in the win



Bernoulli Principle

- 1. Use utility function $u_{ij} = u(e_{ij})$
- 2. Preference $\Phi(a_i)$ is the expected value of the utility values
- 3. "Risk-Utility Function"



Objective Function

$$\max_{i \in [m]} \Phi(a_i) \text{ with } \Phi(a_i) = \sum_{j \in [n]} p_j u(e_{ij})$$

Example

Choose an alternative. Decision matrix:

•
$$\Phi(a_1) = 0 \cdot 0.1 + 3 \cdot 0.5 + 6 \cdot 0.4 = 3.9$$

•
$$\Phi(a_2) = 7 \cdot 0.1 + 5 \cdot 0.5 + 2 \cdot 0.4 = 4.0$$

Assume utility is $u(e) = \sqrt{e}$. Utility matrix:

$$\begin{array}{c|cccc} p & 0.1 & 0.5 & 0.4 \\ \hline a_1 & 0 & 3 & 6 \\ a_2 & 7 & 5 & 2 \\ \end{array}$$



St. Petersburg Paradox

- Utility function circumvents paradox
- For example, for $u(x) = \log_2 x$:

• How is the game assessed?

$$\Phi = \sum_{i=1}^{\infty} p_i u(e_i) = \sum_{i=1}^{\infty} 2^{-i} i = 2$$



Properties

• Assume replacing u(e) with a positive-linear transformation

$$\hat{u}(e) = \alpha + \beta \cdot u(e), \qquad \beta > 0$$

Preference functions:

$$\begin{split} \varPhi(a_i) &= \sum_{j \in [n]} p_j u(e_{ij}) \\ \hat{\varPhi}(a_i) &= \sum_{j \in [n]} p_j \hat{u}(e_{ij}) = \sum_{j \in [n]} p_j \left(\alpha + \beta u(e_{ij})\right) \\ &= \alpha \sum_{j \in [n]} p_j + \beta \sum_{j \in [n]} p_j u(e_{ij}) = \alpha + \beta \varPhi(a_i) \end{split}$$

• $\hat{\varPhi}$ is a positive-linear transformation of \varPhi



Properties

- Preference functions can be positively linearly transformed without changing the decision
- Arbitrariness of the zero point and unit of measurement
- Sign of Φ is not an indicator of whether the result is pleasant or unpleasant

FAU M. Hartisch Decision Theory June 24, 2024 52/61



Determination of Φ

- Comparison between hypothetical choices of:
 - Certain income
 - Simple chance

Simple Chance

- Exactly two possible outcomes e' and e''
- Probabilities p' and p'' = 1 p'
- Abbreviated as (e'; p'; e")
- e' is the favorable outcome
- p': Probability of success



Example

- Lottery:
 - o 99% lose
 - 1% win 1000 euros
- e' = 1000, p' = 0.01
- e'' = 0, p'' = 0.99
- Notation: (1000; 0.01; 0)



Utility Elicitation – First Step

- All possible outcomes e_{ij} should lie in the range [e'', e']
- Assign arbitrary utilities 1 = u' = u(e') > u'' = u(e'') = 0

Example

- All outcomes lie in the range between 0 and 100
- Define:

$$\circ$$
 $e' = 100, u' = u(100) = 1$

$$e'' = 0, u'' = u(0) = 0$$



Utility Elicitation – Second Step

- Decision maker is faced with the choice between
 - \circ A certain alternative with outcome \overline{e}
 - \circ A simple chance (e'; p'; e'')

Example

- Let's assume choosing $\overline{e} = 25$
- Compare between:
 - Certain outcome of 25
 - \circ Simple chance (100; p'; 0)

FAU M. Hartisch Decision Theory



Utility Elicitation – Third Step

- p' is determined so that both alternatives are considered equally good
- Critical value: $p^*(\overline{e})$

Example

- p' close to 1: (100; p'; 0) better than $\overline{e} = 25$
- p' close to 0: (100; p'; 0) worse than $\overline{e} = 25$
- By iterative questioning, $p^*(25)$ is determined
- Let's assume $p^*(25) = 0.5$



Utility Elicitation – Fourth Step

- \overline{e} is as good as $(e'; p^*(\overline{e}); e'')$
- Therefore,

$$u(\overline{e}) = p^*(\overline{e})u(e') + (1 - p^*(\overline{e}))u(e'')$$

• Through normalization: $u(\overline{e}) = p^*(\overline{e})$

Example

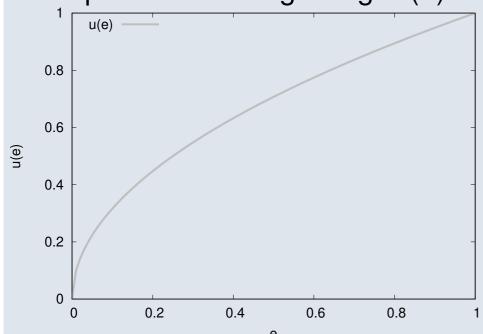
$$u(25) = 0.5u(100) + (1 - 0.5)u(0) = 0.5 \cdot 1 = 0.5$$

FAU M. Hartisch Decision Theory



Course of the Risk-Utility Function

Example from the beginning: $u(e) = \sqrt{e}$



- Larger outcome yields greater utility: $\frac{du}{de} > 0$
- What about the curvature?
 - $\circ \frac{d^2u}{de^2} < 0$: concave increasing
 - $\circ \frac{d^2u}{de^2} > 0$: convex increasing
 - $\circ \frac{d^2u}{de^2} = 0$: linear increasing



Course of the Risk-Utility Function

- Concave course:
 - Certainty equivalent less than expected value
 - Risk-averse
- Convex course:
 - Certainty equivalent greater than expected value
 - Risk-seeking
- Linear course:
 - Certainty equivalent equal to expected value
 - Risk-neutral

FAU M. Hartisch Decision Theory



Quiz

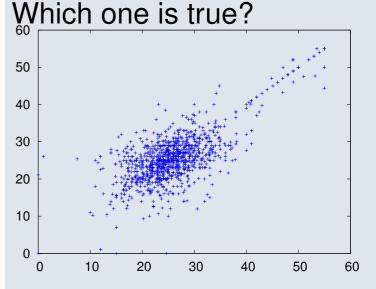
Question 1

What combination (x_1, x_2) yields a portfolio with the minimum variance?

Question 2

Let u be a linear utility function with u(8) = 0 and u(18) = 1. What is the expected utility?

- 1. $\Phi(a_1) = 0.20$
- **2**. $\Phi(a_1) = 0.32$
- 3. $\Phi(a_1) = 0.41$



- 1. positive correlation
- 2. no correlation
- 3. negative correlation



Quiz

Question 1

What combination (x_1, x_2) yields a portfolio with the minimum variance?

Solution

For $x_1 = 1/3$ and $x_2 = 2/3$, all values are zero \rightarrow variance is zero.

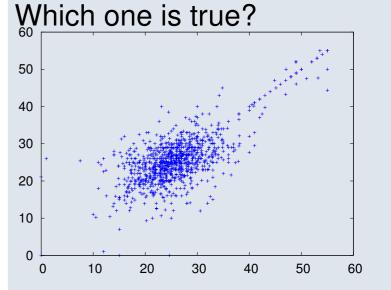
Question 2

Let u be a linear utility function with u(8) = 0 and u(18) = 1. What is the expected utility?

1.
$$\Phi(a_1) = 0.20$$

2.
$$\Phi(a_1) = 0.32$$

3.
$$\Phi(a_1) = 0.41$$



- 1. positive correlation
- 2. no correlation
- 3. negative correlation



Quiz

Question 1

What combination (x_1, x_2) yields a portfolio with the minimum variance?

Solution

For $x_1 = 1/3$ and $x_2 = 2/3$, all values are zero \rightarrow variance is zero.

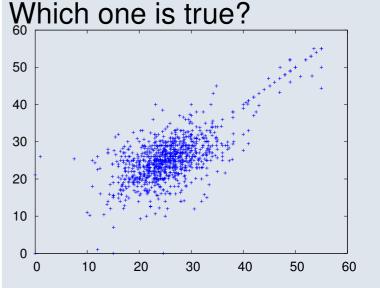
Question 2

Let u be a linear utility function with u(8) = 0 and u(18) = 1. What is the expected utility?

1.
$$\Phi(a_1) = 0.20$$

2.
$$\Phi(a_1) = 0.32$$

3.
$$\Phi(a_1) = 0.41$$



- 1. positive correlation
- 2. no correlation
- 3. negative correlation



Quiz

Question 1

What combination (x_1, x_2) yields a portfolio with the minimum variance?

Solution

For $x_1 = 1/3$ and $x_2 = 2/3$, all values are zero \rightarrow variance is zero.

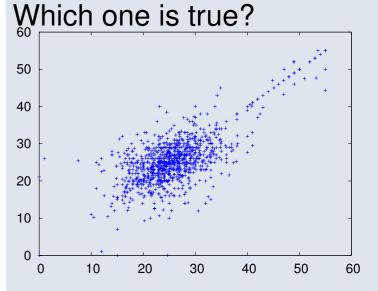
Question 2

Let u be a linear utility function with u(8) = 0 and u(18) = 1. What is the expected utility?

1.
$$\Phi(a_1) = 0.20$$

2.
$$\Phi(a_1) = 0.32$$

3.
$$\Phi(a_1) = 0.41$$



- 1. positive correlation
- 2. no correlation
- 3. negative correlation