MIDTERM EXAM

YOU CAN USE THE VERSO OF YOUR EXAM AS A DRAFT OR AS ADDITIONAL SPACE TO ANSWER QUESTIONS. STATE CLEARLY HOW YOU ARE USING IT.

PROBLEM 1 (28 points)

Let's consider a problem encountered by a company that must choose amongst 4 alternatives for the production of a product which <u>costs</u> are uncertain. The 3 scenarios under consideration form the set of situations the company contemplates and form the set of all elementary events of the problem.

Table of costs obtained depending on the chosen alternative and realized scenario

| | Scenarios (probabilities) | | | |
|--------------|---------------------------|----------|----------|--|
| Alternatives | S1 (75%) | S2 (10%) | S3 (15%) | |
| XA | 1500 | 0 | 1600 | |
| XВ | 750 | 2000 | 1100 | |
| XC | 1100 | 1500 | 0 | |
| XD | 1100 | 1500 | 100 | |

- a) Identify, if possible, an alternative for which the cost is stochastically dominated, but not deterministically dominated, by the cost generated by another alternative. (4 points)
- b) Identify, if possible, an alternative for which the cost is deterministically dominated by the cost generated by another alternative. (2 points)
- c) Calculate the following statistics for the cost generated by alternatives xA. (14 points)

| Statistic | Values for xA |
|---|---------------|
| Expected value | |
| Standard deviation | |
| Probability that no cost is incurred | |
| Value at risk for a confidence level of 85% | |
| Conditional value at risk for a confidence level of 85% | |
| Value at risk for a confidence level of 70% | |
| Conditional value at risk for a confidence level of 70% | |

d) Describe how one could generate a scenario among S1, S2, and S3 randomly based on a random variable that is uniformly distributed on the interval [0, 1]. The probability of obtaining S1, S2, or S3 should respectively be 75%, 10%, and 15%. **(4 points)**

e) Based on the following table, compare the risks associated to the costs generated by alternatives xB and xC. (4 points)

| Statistic | Values for xB | Values for xC |
|---|---------------|---------------|
| Expected value | \$927,50 | \$975 |
| Standard deviation | \$378,31 | \$426,47 |
| Probability that no cost is incurred | 0% | 15% |
| Value at risk for a confidence level of 90% | \$1100 | \$1100 |
| Conditional value at risk for a confidence level of 90% | \$2000 | \$1500 |

a) This is a description of the cumulative distribution functions of the cost achieved by the 4 alternatives.

| | 0 | 100 | 750 | 1100 | 1500 | 1600 | 2000 |
|----|-----|-----|-----|------|------|------|------|
| xA | 10% | 10% | 10% | 10% | 85% | 100% | 100% |
| хB | 0% | 0% | 75% | 90% | 90% | 90% | 100% |
| хC | 15% | 15% | 15% | 90% | 100% | 100% | 100% |
| xD | 0% | 15% | 15% | 90% | 100% | 100% | 100% |

We observe that the cumulative distribution function of xA is always below of the distribution function of xC. It is therefore the case that for all level of cost xA has always less chances to be below this level than xC. xC therefore dominates stochastically xA. But, xC does not dominate deterministically xA since for the scenario S2 it leads to a greater financial loss than xA (1500 for xC, 0 for xA).

xC stochastically dominates xD is not adequate since it also deterministically dominates xA.

- **b)** The alternative xC deterministically dominates the alternative xD since it always leads to a lower cost regardless of which scenario happens.
- c) We can do the following calculations:

For xA:

Expected value = 75%*(1500)+15%*(1600)+10%*(0) = 1365Standard deviation =

$$\sqrt{75\% * (1500 - 1365)^2 + 15\% * (1600 - 1365)^2 + 10\% * (0 - 1365)^2} = 456,37$$

No cost probability = 10%

85% Value at risk = 1500 since we have more than 85% chances of having a cost lower than 1500\$.

85% CVaR = 1600\$ since in the worst 15% scenarios we achieve on average cost of (15%/15%)*(1600)=1600\$.

70% Value at risk = 1500 since we have more than 70% chances of having a cost lower than 1500\$.

70% CVaR = 1550\$ since in the worst 30% scenarios we achieve on average cost of (15%/30%)*(1500)+(15%/30%)*(1600)=1550\$.

| Statistic | Values for xA | |
|--|---------------|--|
| Expected value | 1365 | |
| Standard deviation | 456,37 | |
| Probability of financial loss | 10% | |
| Value at risk for a confidence level of 85 % | 1500 | |

| Conditional value at risk for a confidence level of 85 % | 1600 |
|--|------|
| Value at risk for a confidence level of 70 % | 1500 |
| Conditional value at risk for a confidence level of 70 % | 1550 |

d) This is obtained by following the rule below:

| If U <= 75% | Conclude S1 |
|--|-------------|
| If 75% <u<=85%< th=""><th>Conclude S2</th></u<=85%<> | Conclude S2 |
| If U>85% | Conclude S3 |

e) The expected value of the cost of xB is lower than for xC and its standard deviation is lower. On the other hand, we see that the probability of no cost is larger for xC and the 90%-CVaR is much higher for xB than for xC. This indicates that the risks are larger for xB than xC. The best alternative here depends on the decision maker's aversion to risk. Is he willing to take the risk of a larger cost in order to get lower cost on average?

PROBLEM 2 (22 points)

A company must decide on the construction of a production facility. The company may decide to build a large or a small facility; in the latter case, the company may also decide to increase the capacity of the facility after two years of operation. Three possible scenarios are considered for the demand of the products it will produce in the facility:

- Scenario #1: Low demand for 10 years
- Scenario #2: <u>High demand</u> for the first 2 years, <u>low demand</u> for the next 8 years
- Scenario #3: High demand for 10 years

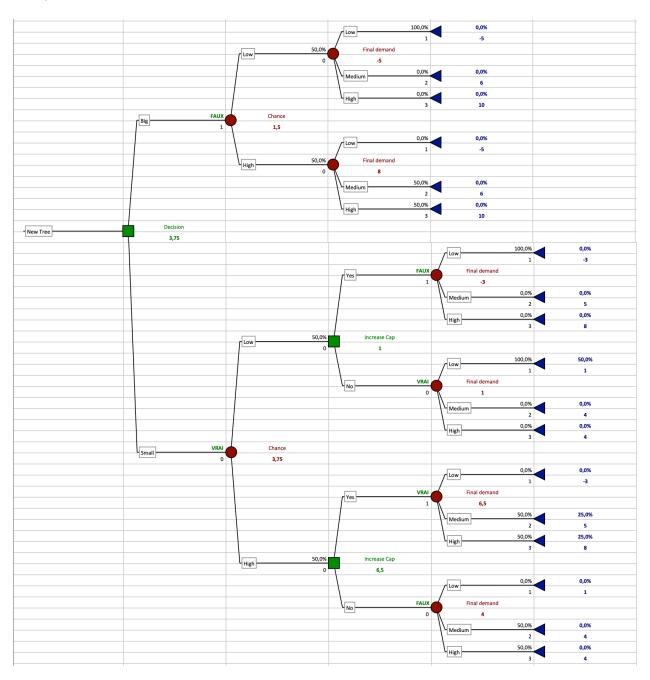
The table below summarizes what is the net present value of the cash flow generated by the construction of each type of facility under each type of demand scenario and the probabilities of each of these scenarios.

Net present value (in millions of dollars)

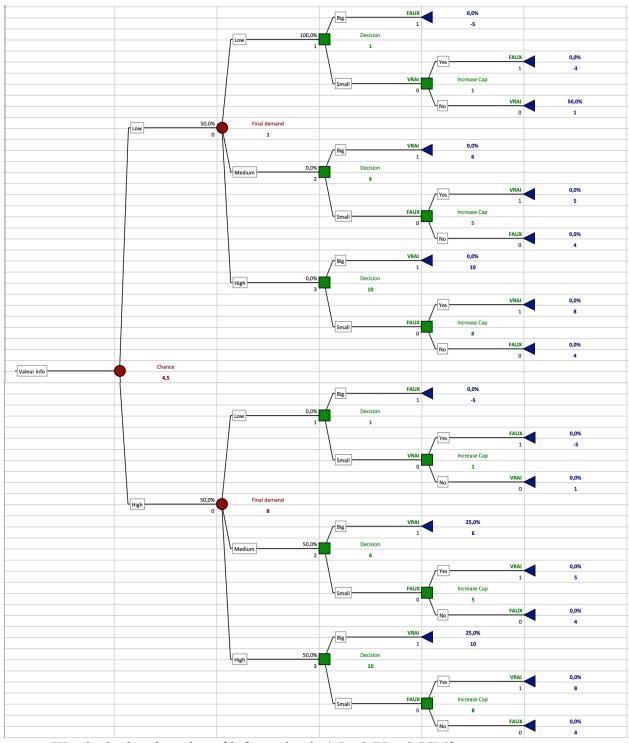
| | Scenario #1 (50%) | Scenario #2 (25%) | Scenario #3 (25%) |
|---|----------------------|----------------------|----------------------|
| Big facility | -5 | 6 | 10 |
| Small facility | 1 | 4 | 4 |
| Small facility with increased capacity after 2 nd year | -3 | 5 | 8 |

- a) Draw and solve a decision tree that allows to identify whether this company should build a large or small facility and whether (and under which conditions) it should increase the capacity of the small facility. What is the optimal strategy for this company? You can assume that the company considers the expected net present value as the only measure of success for this project. (10 points)
- b) Consider now that the company has the opportunity to learn exactly what the demand for the products will be in the next 10 years before deciding which facility to construct. What would be the largest amount of money that this company would be willing to pay to obtain this information? (6 points)
- c) Consider now that the company is instead able to hire the services of a consulting firm that will predict only what the demand will be in the next 2 years. What would be the largest amount of money that this company would be willing to pay to obtain this information? (6 points)

a) Here is the tree and its resolution.

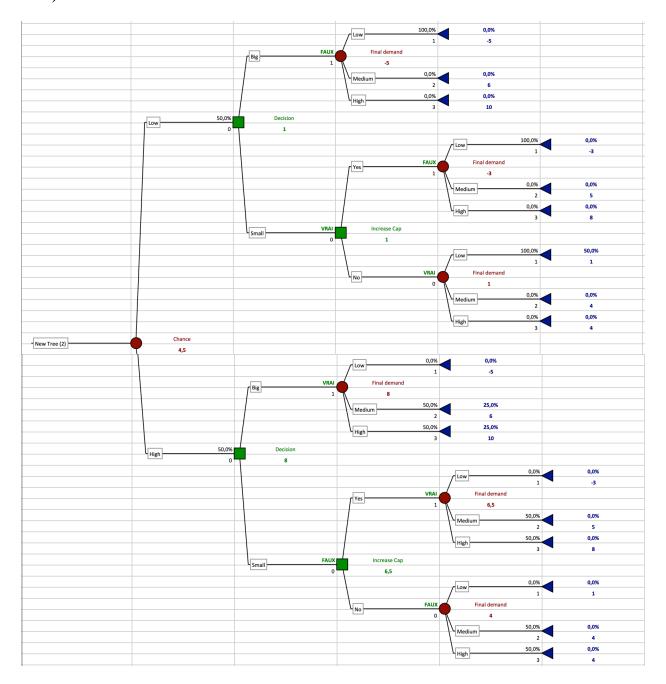


b) Here is the tree and its resolution.



We obtain that the value of information is 4.5 - 3.75 = 0.75M\$.

c) Here is the tree and its resolution.

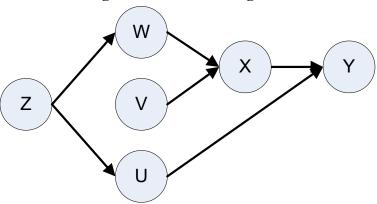


We obtain that the value of information is again 4,5-3,75 = 0,75M\$.

PROBLEM 3 (26 points)

We consider the following influence diagram involving the random variables U, V, W, X, Y and Z and satisfying the Markov property presented in class:

Figure 1: Influence Diagram



a) Identify the 5 independences that can be deduced from the above network's structure amongst the following 10 proposals: (5 points)

| Independence | Can be deduced from the diagram? (Yes / No) | Independence | Can be deduced from the diagram? (Yes / No) |
|----------------------|---|------------------------|---|
| V ™ W | | ZΨV | |
| V [⊥] W Y | | Z╨Y | |
| X [⊥] U Y | | V [⊥] L W Z | |
| X [⊥] U Z | | Z [⊥] Y U | |
| ZΨW | | Y [⊥] Z U,X | |

For example, "Y $\perp \!\!\! \perp Z \mid U,X$ " describes the fact that Y is independent of Z if U and X are known.

- b) Enumerate the distribution functions that must be defined in order to obtain a complete and compact model (i.e., that exploits independences) of this problem's uncertainty. (3 points)
- c) If all variables are Bernoulli random variables (with realizations 0 or 1) except for the variables V and W that possess four possible realizations (0, 1, 2, or 3), how many parameters must be estimated in total to define these distributions? (3 points)
- d) If all variables are Bernoulli random variables (with realizations 0 or 1) except for the variables V and W that possess four possible realizations (0, 1, 2, or 3), how many parameters must be estimated in total to define these distributions in the event that no independence assumption could be made in this problem? (3 points)

e) Considering that you have identified the following distribution models for realizations of the pair (Z, W), use Bayes' theorem to reformulate this model in a way that can be used in the **Influence Diagram** presented in **Figure 1**. (6 points)

| Value for W | Probability |
|-------------|-------------|
| (w) | P(W=w) |
| 0 | 18% |
| 1 | 22% |
| 2 | 40% |
| 3 | 20% |

| Value for W | Value for Z | Probability |
|-------------|-------------|--------------|
| (w) | (z) | P(Z=z W=w) |
| 0 | 0 | 20% |
| 0 | 1 | 80% |
| 1 | 0 | 50% |
| 1 | 1 | 50% |
| 2 | 0 | 50% |
| 2 | 1 | 50% |
| 3 | 0 | 80% |
| 3 | 1 | 20% |

f) You are now considering characterising a conditional probability model for variable U when Z is known. You would actually like the distribution of U when Z is known to follow a Poisson distribution. Describe the model that would be obtained using the following data (each entry being obtained independently from each other) when employing a Bayesian approach. (6 points)

| Index of the observation | Observed value for Z | Observed value for U | Index of the observation | Observed value for Z | Observed value for U |
|--------------------------|----------------------|----------------------|--------------------------|----------------------|----------------------|
| 1 | 0 | 6 | 9 | 1 | 1 |
| 2 | 0 | 17 | 10 | 1 | 2 |
| 3 | 0 | 14 | 11 | 1 | 1 |
| 4 | 0 | 14 | 12 | 1 | 3 |
| 5 | 0 | 11 | 13 | 1 | 5 |
| 6 | 0 | 14 | 14 | 1 | 4 |
| 7 | 0 | 11 | 15 | 1 | 3 |
| 8 | 0 | 6 | 16 | 1 | 1 |

a)

| Independence | Can be deduced from the diagram? (Yes / No) | Independence | Can be deduced from the diagram? (Yes / No) | |
|------------------------|---|------------------------|---|--|
| V | Yes | ZΨV | Yes | |
| V [⊥] W Y | No | Z╨Y | No | |
| X [⊥] L U Y | No | V | Yes | |
| X [⊥] U Z | Yes | Z # Y U | No | |
| Z H W | No | Y [⊥] Z U,X | Yes | |

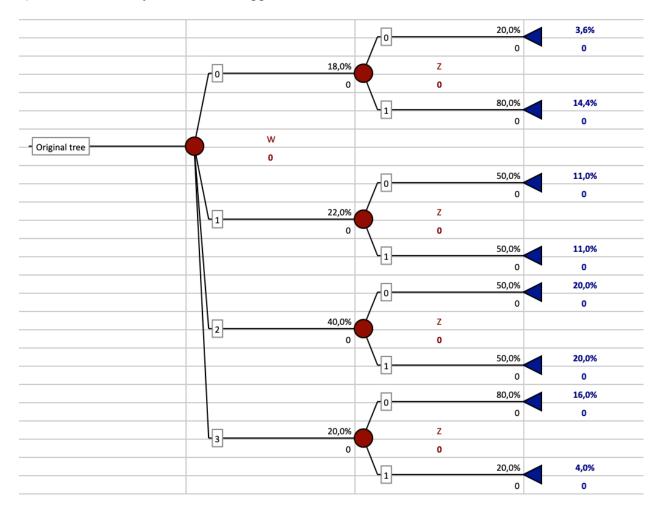
- **b)** The following distributions must be defined : P(Z), P(V), P(W|Z), P(U|Z), P(X|W,V), P(Y|X,U)
- **c)** The following probabilities must be estimated:

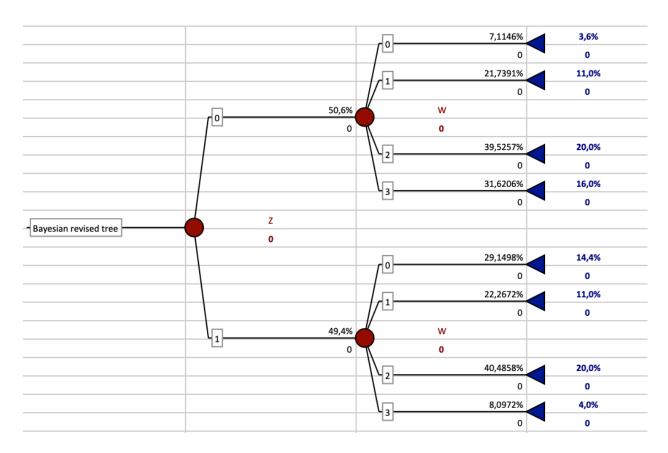
$$\begin{split} &P(Z=1)\\ &P(V=1), P(V=2), P(V=3)\\ &P(W=1|Z=0), P(W=1|Z=1),\\ &P(W=2|Z=0), P(W=2|Z=1),\\ &P(W=3|Z=0), P(W=3|Z=1),\\ &P(U=1|Z=0), P(U=1|Z=1)\\ &P(X=1|W=0,V=0), P(X=1|W=0,V=1), P(X=1|W=0,V=2), P(X=1|W=0,V=3),\\ &P(X=1|W=1,V=0), P(X=1|W=1,V=1), P(X=1|W=1,V=2), P(X=1|W=1,V=3),\\ &P(X=1|W=2,V=0), P(X=1|W=2,V=1), P(X=1|W=2,V=2), P(X=1|W=2,V=3),\\ &P(X=1|W=3,V=0), P(X=1|W=3,V=1), P(X=1|W=3,V=2), P(X=1|W=3,V=3),\\ &P(Y=1|X=0,U=0), P(Y=1|X=0,U=1), P(Y=1|X=1,U=0), P(Y=1|X=1,U=1)\\ &In total, 1+3+6+2+16+4=32 probabilities \end{split}$$

d) The probabilities of every combinations of events for the variables U, V, W, X, Y & Z must be estimated

The last probability of P(U=1,V=3,W=3,X=1,Y=1,Z=1) can be deducted from the other by determining what is missing to make 100%. There is therefore 2*4*4*2*2*2-1=255 probabilities to estimate.

e) Here is how Baye's theorem is applied.





f) If we apply the Bayesian approach, then we need to consider the Poisson parameter λ_0 for the case where Z=0 to be different than the Poisson parameter λ_1 when Z=1. In particular, we initially assume that both of them is distributed according to the conjugate prior of the Poisson distribution (i.e., the Gamma distribution). Hence, λ_0 is distributed according to Gamma(k_0 , θ_0) while λ_1 is distributed according to Gamma(k_1 , θ_1). Once, the data in the table is observed we can update the choice of distribution for the two λ 's. For the case Z=0, we have $k'_0 = k_0 + 93$, $\theta'_0 = \theta_0/(8\theta_0 + 1)$. In the case of Z=1, we have rather that $k'_1 = k_1 + 20$ and $\theta'_1 = \theta_1/(8\theta_1 + 1)$.

PROBLEM 4 (24 points)

Let's consider a similar company as discussed in problem 2. This time the company can only choose between building a large or a small facility. The company has run a Monte-Carlo simulation in order to compare the risks of the two alternatives. Doing so it obtained 40 samples for each alternative, each of which should be considered an independent and identically distributed realization from the distribution of net present value (NPV) achieved by each facility type. These samples are presented in the following two tables in the order from the smallest to the largest simulated scenario. (Note that the units of NPV are in thousands of dollars, i.e. K\$)

| Ordered samples obtained from Monte-Carlo simulation of <u>Large Facility</u> | | | | | | | |
|---|-------------------|------|-------------------|------|-------------------|------|-------------------|
| Rank | Value (in K\$) | Rank | Value (in K\$) | Rank | Value (in K\$) | Rank | Value (in K\$) |
| 1 | -153 | 11 | 54 | 21 | 119 | 31 | 149 |
| 2 | -63 | 12 | 59 | 22 | 123 | 32 | 151 |
| 3 | -52 | 13 | 70 | 23 | 126 | 33 | 156 |
| 4 | -50 | 14 | 87 | 24 | 131 | 34 | 157 |
| 5 | -47 | 15 | 91 | 25 | 137 | 35 | 162 |
| 6 | -5 | 16 | 93 | 26 | 140 | 36 | 163 |
| 7 | 3 | 17 | 108 | 27 | 141 | 37 | 171 |
| 8 | 35 | 18 | 110 | 28 | 145 | 38 | 172 |
| 9 | 40 | 19 | 113 | 29 | 146 | 39 | 176 |
| 10 | 51 | 20 | 113 | 30 | 147 | 40 | 181 |

| Ordered samples obtained from Monte-Carlo simulation of Small Facility | | | | | | | |
|---|-------------------|------|-------------------|------|-------------------|------|-------------------|
| Rank | Value (in K\$) | Rank | Value (in K\$) | Rank | Value (in K\$) | Rank | Value (in K\$) |
| 1 | -42 | 11 | 17 | 21 | 47 | 31 | 84 |
| 2 | -31 | 12 | 24 | 22 | 48 | 32 | 85 |
| 3 | -23 | 13 | 28 | 23 | 56 | 33 | 94 |
| 4 | -23 | 14 | 33 | 24 | 58 | 34 | 95 |
| 5 | -10 | 15 | 34 | 25 | 63 | 35 | 106 |
| 6 | -4 | 16 | 35 | 26 | 66 | 36 | 109 |
| 7 | 2 | 17 | 35 | 27 | 70 | 37 | 122 |
| 8 | 3 | 18 | 40 | 28 | 72 | 38 | 123 |
| 9 | 11 | 19 | 41 | 29 | 76 | 39 | 130 |
| 10 | 14 | 20 | 47 | 30 | 79 | 40 | 145 |

According to this sample set, the expected net present value of the large facility is \$91,25K and its standard deviation is \$79,44K, while for the small facility the mean is \$48,98K and the standard deviation is \$46,62K.

- a) Calculate a 86% confidence interval for the expected value of the net present value of the <u>large</u> <u>facility</u>. (You can use the table provided in the **ANNEX**.) (4 points)
- b) How many samples would be needed in order for the 86% confidence interval calculated in a) to have a size that is smaller than \$0,1K? (You can use the table provided in the ANNEX.) (2 points)
- c) Estimate the probability that the net present value for the <u>large facility</u> is strictly lower than \$0K. Considering that this probability is in fact the expected value of the random variable defined as

$$Z = \begin{cases} 1 & if the NPV is < 0\$\\ 0 & if the NPV is \ge 0\$ \end{cases}$$

determine a 86% confidence interval for this probability. (You can use the table provided in the **ANNEX**.) **(6 points)**

- **d)** Estimate the 20th percentile of the net present value of the <u>large facility</u> and an 86% confidence interval for this percentile. (You can use the table provided in the **ANNEX**.) **(6 points)**
- e) Based on the results of this Monte-Carlo simulation, what would be the best type of facility to build? Present your conclusion by discussing the compromise between the two alternatives in terms of risk and returns. (6 points)

- a) We can compute this answer using the formula seen in class. The inverse normal distribution evaluated at 0.93=1-((1-0.86)/2) is equal 1,4758. The confidence interval is therefore equal to : $\left[91,25-1,4758\cdot\frac{79,44}{\sqrt{40}}\right]=\left[72,71,\ 109,79\right]$.
- **b)** The interval width can be estimated as $2 \cdot 1,4758 \cdot \frac{79,44}{\sqrt{M}}$, where M is the number of samples. To obtain an interval of width less than $0,1, M \ge 5\,497\,857$ samples are required
- c) We can estimate this probability by computing the frequency in the sample set. This tells us that the probability is $\frac{6}{40} = 15\%$. Considering that this probability is the expected value of Z, we can apply the formula for confidence interval of expected value. In the data, Z takes 6 times the value of 1 and 34 times the value of 0. We can estimate its standard deviation as $\sqrt{\left(\frac{6}{40}\right) \cdot (1 0.15)^2 + \left(\frac{34}{40}\right) \cdot (0 0.15)^2} = 0.3571$. The confidence interval is therefore: $[0.15 1.4758 * 0.3571 / 40^{0.5}, 0.15 + 1.4758 * 0.3571 / 40^{0.5}] = [0.0667, 0.2333]$. Namely, the probability should be between 6.67% and 23,33%.
- d) We can estimate the 20^{th} percentile using the 8^{th} value in the ordered list (i.e. 40 * 0.2 = 8). We therefore estimate this percentile at 35 K\$.

In the case of the confidence interval, we use the formula studied in class. First, we need to compute the Δ .

$$\Delta = 1,4758 \sqrt{0,2 \times (1-0,2)} / \sqrt{40} = 0,1044$$

Then, we obtain that the index for the left value is $40 \cdot (0.2 - 0.1044) = 3.82 \rightarrow 3$ while the index for the right value is $40 \cdot (0.2 + 0.1044) = 12.17 \rightarrow 13$, to which we add one as in the slides. Overall, we get that the 20^{th} percentile is between [-52, 87] with a probability of 86%.

e) The small facility has a lower expected NPV but also smaller standard deviation. This indicates that although it is the less opportunistic, it is interesting in the way it protects against risk. In particular, one can estimate the 80%-CVaR of this facility to be 16K\$ while the 80%-CVaR of the larger facility is 41,5K\$. This is due to the worst-case scenario of -153 K\$ for the large facility compared to -42 K\$ for the small one.

On the other hand, the large facility achieves almost twice more expected NPV and is therefore quite attractive for a company that does not worry so much about risk. One should be careful however in terms of how much more expected NPV is reached with this alternative given that the true NPV could be as low as 72 K\$ compared to an estimate of 49K\$ for the small facility. Perhaps, one should perform a more extensive simulation to estimate this expected NPV more precisely.

ANNEX

| Inverse distribution function for a normal distribution (mean = 0, standard deviation = 1) | | | | | | | |
|--|---------------------|-------------|---------------------|-------------|---------------------|-------------|---------------------|
| Probability | F ⁻¹ (p) | Probability | F ⁻¹ (p) | Probability | F ⁻¹ (p) | Probability | F ⁻¹ (p) |
| 0,01 | -2,32634787 | 0,26 | -0,64334541 | 0,51 | 0,02506891 | 0,76 | 0,70630256 |
| 0,02 | -2,05374891 | 0,27 | -0,61281299 | 0,52 | 0,05015358 | 0,77 | 0,73884685 |
| 0,03 | -1,88079361 | 0,28 | -0,58284151 | 0,53 | 0,07526986 | 0,78 | 0,77219321 |
| 0,04 | -1,75068607 | 0,29 | -0,55338472 | 0,54 | 0,10043372 | 0,79 | 0,80642125 |
| 0,05 | -1,64485363 | 0,30 | -0,52440051 | 0,55 | 0,12566135 | 0,80 | 0,84162123 |
| 0,06 | -1,55477359 | 0,31 | -0,49585035 | 0,56 | 0,15096922 | 0,81 | 0,8778963 |
| 0,07 | -1,47579103 | 0,32 | -0,4676988 | 0,57 | 0,17637416 | 0,82 | 0,91536509 |
| 0,08 | -1,40507156 | 0,33 | -0,43991317 | 0,58 | 0,20189348 | 0,83 | 0,95416525 |
| 0,09 | -1,34075503 | 0,34 | -0,41246313 | 0,59 | 0,22754498 | 0,84 | 0,99445788 |
| 0,10 | -1,28155157 | 0,35 | -0,38532047 | 0,60 | 0,2533471 | 0,85 | 1,03643339 |
| 0,11 | -1,22652812 | 0,36 | -0,35845879 | 0,61 | 0,27931903 | 0,86 | 1,08031934 |
| 0,12 | -1,17498679 | 0,37 | -0,33185335 | 0,62 | 0,30548079 | 0,87 | 1,12639113 |
| 0,13 | -1,12639113 | 0,38 | -0,30548079 | 0,63 | 0,33185335 | 0,88 | 1,17498679 |
| 0,14 | -1,08031934 | 0,39 | -0,27931903 | 0,64 | 0,35845879 | 0,89 | 1,22652812 |
| 0,15 | -1,03643339 | 0,40 | -0,2533471 | 0,65 | 0,38532047 | 0,90 | 1,28155157 |
| 0,16 | -0,99445788 | 0,41 | -0,22754498 | 0,66 | 0,41246313 | 0,91 | 1,34075503 |
| 0,17 | -0,95416525 | 0,42 | -0,20189348 | 0,67 | 0,43991317 | 0,92 | 1,40507156 |
| 0,18 | -0,91536509 | 0,43 | -0,17637416 | 0,68 | 0,4676988 | 0,93 | 1,47579103 |
| 0,19 | -0,8778963 | 0,44 | -0,15096922 | 0,69 | 0,49585035 | 0,94 | 1,55477359 |
| 0,20 | -0,84162123 | 0,45 | -0,12566135 | 0,70 | 0,52440051 | 0,95 | 1,64485363 |
| 0,21 | -0,80642125 | 0,46 | -0,10043372 | 0,71 | 0,55338472 | 0,96 | 1,75068607 |
| 0,22 | -0,77219321 | 0,47 | -0,07526986 | 0,72 | 0,58284151 | 0,97 | 1,88079361 |
| 0,23 | -0,73884685 | 0,48 | -0,05015358 | 0,73 | 0,61281299 | 0,98 | 2,05374891 |
| 0,24 | -0,70630256 | 0,49 | -0,02506891 | 0,74 | 0,64334541 | 0,99 | 2,32634787 |
| 0,25 | -0,67448975 | 0,50 | 0 | 0,75 | 0,67448975 | 1,00 | ∞ |