Mean-Variance Space for Evaluations

Leo Schubert,[1] Constance University of Applied Sciences

Abstract

The effect on the mean-variance space of restrictions on a variable is investigated in this paper. A restriction may be the placing of upper and lower bounds on a variable. Another limitation is the loss of the continuity of a variable.
Average marks for examinations are considered in an application of this limited mean-variance space. In this case, the bounds are given by the highest and the lowest possible mark (e.g. 1.0 and 5.0). The limitation of the mean-variance space depends on the number of students who participate in the examination. The restriction of the loss of continuity is shown by the use of discrete marks (e.g. 1.0, 1.3, 1.7, 2.0, …). Furthermore, the Target-Shortfall-Probability lines are integrated into the mean-variance space. These lines are used to indicate the proportion of students who have good or very good marks in the examination. In financial markets, Target-Shortfall-Probability is used as a risk criterion.

Key Words: Dispersion, Mean-Variance Space, Target-Shortfall-Probability, Inflation of best grade points (JEL Classification System: A20, C10, C46, I21).

1. Introduction

The aim of most students is to leave university with very good examination results. In the past two decades the percentage of students with excellent marks has risen[2]. High tuition fees charged by universities have their own influence on this development, and seem to be part of modern society.[3] Today, the problem of the inflation of marks in examinations can also be observed in universities without tuition fees. There may be different factors affecting this problem. To study the inflation for different courses and examinations in a mean-variance space, this paper investigates the limitation of the space. The first restriction is based on the range of marks, which may be from 1.0 for the best exam paper to 5.0 for the worst. The second restriction occurs when the scale of marks is not continuous, such as when the marks may only have the following values: 1.0, 1.3, 1.7, 2.0, 2.3, 2.7, 3.0, 3.3, 3.7, 4.0, 4.3, 4.7, 5.0. Both restrictions have an impact on the mean-variance space of the average marks in examinations.
The mean-variance space with its limitations gives a framework for depicting the average marks for courses, and additionally offers the possibility of integrating the so-called “Target-Shortfall-Probability” lines (TSP lines).[4] The position of these lines in the mean-variance space gives information about the proportion of students in an examination who have good or very good marks. The TSP is a risk criterion in financial markets and portfolio optimization.
The effect of the first limitation will be discussed in chapters 3 and 4, while chapters 5 and 6 are dedicated to the TSP with respect to the second limitation of discrete marks.
Although the variance of marks is observed and depicted, this paper does not recommend maximizing this criterion. Maximization would produce examinations where the only marks were 1.0 and 5.0. This could not be a desirable aim.

Other measures of dispersion were not tested. The absolute deviation is less usual, and the recommendation would be to use the median of the marks instead of the mean. Furthermore, the integration of the TSP lines would not be possible. Coefficients of concentration[5] and indexes to measure biodiversity[6] were not used to measure the dispersion in the marks. These coefficients often

---

[1] Schubert@HTWG-Konstanz.de

[2] Preuss R.: Zu gute Noten an deutschen Hochschulen, SZ, 11.11.2012, p. 1.

[3] The problem of the inflation of the best grade points at universities in the USA seems to be caused by the character of American society (see Wallace, D. F., (2009), p. 145).

[4] See e.g. Schubert L. (2002).

[5] See e.g. Bamberg, G., Baur, F., Krapp M., (2012), pp. 22ff.

[6] See e.g. Bouza Herrera, C., Schubert, L. (2003).

measure the deviation of equally distributed results, which is not a realistic assumption for the dispersion of marks.

2. Applications of the mean-variance space

The mean-variance space is well known in capital market theory. If portfolios of assets are to be efficient, the variance of these portfolios has to be minimized. The variance depends on the variance of the return on the single assets and also on the covariance between two assets.
In this paper, the variance with respect to the standard deviation of the marks in an examination is not minimized. Instead, the maximal possible variance of the marks will be sought. Additionally, the marks have neither variance nor covariance.
Figure 2-1 shows, for two assets A and B (with mean returns of 1% and 5%, equal standard deviation s=2 and covariance cov_{AB}=0), the efficient portfolios, which are on the “up” side of the minimal variance point (MVP).[7] To illustrate the difference between this and the example used in this paper, the lowest mark is selected as 1.0 and the highest as 5.0. The curve of the maximization of the variance that is now depicted is a reverse of the curve of the portfolio example. The different forms of this curve will be derived in the following chapter.

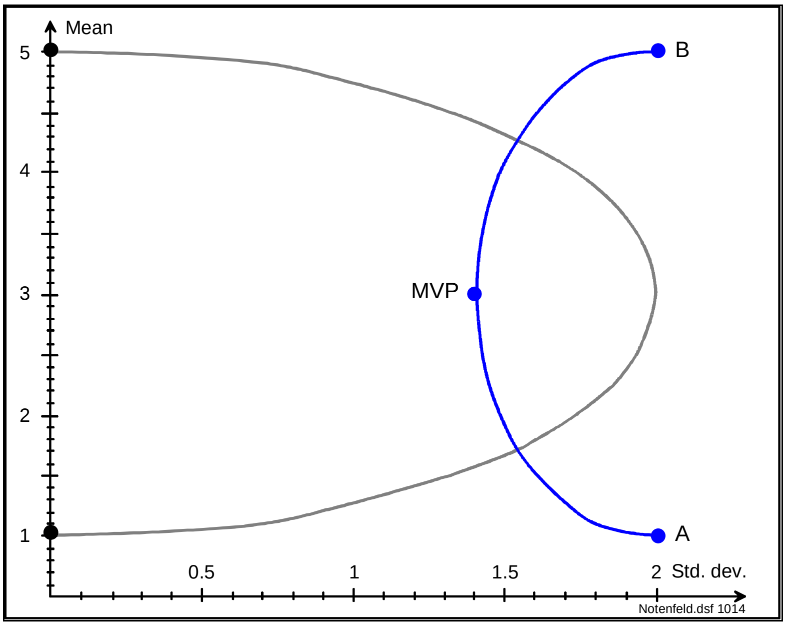

Figure 2-1: Mean–Standard deviation space applied in the selection of portfolios and in the illustration of the average marks in an examination.

3. The maximal standard deviation with marks from 1.0 to 5.0

The maximal standard deviation for marks between 1.0 and 5.0 depends on the number n of participants in the examination. Therefore, the maximal standard deviation will be shown first with the mean $\bar{x} = 3.0$ for an even and an odd number of participants, and then it will be shown in a more general way.

For <u>an even number</u> n of participants:
The maximal standard deviation always occurs when n/2 students score $x_i = 1.0$ (i=1, …, n/2) and all the other entrants score $x_i = 5.0$ (for i=n/2+1, …, n). The mean of the marks will in this case be:

$$\bar{x} = \frac{1}{n} \cdot \left(\frac{n}{2} \cdot 1.0 + \frac{n}{2} \cdot 5.0 \right) = \frac{1}{n}(3 \cdot n) = 3.0 \,. \qquad (3\text{-}1)$$

[7] Computation of the efficient frontier: see Grundmann, W., Luderer, B., (2003), pp. 146f.

For this mean, the variance is:

$$s^2(\overline{x}=3)=\frac{1}{n}\cdot\left(\frac{n}{2}\cdot(1.0-3.0)^2+\frac{n}{2}\cdot(5.0-3.0)^2\right)=\frac{1}{n}(4\cdot n)=4.0 \qquad (3\text{-}2)$$

and the standard deviation $s=2$ (see Figure 2-1).

For <u>an odd number n</u> of participants:
In this case, (n-1)/2 students get 1.0 and the same number get 5.0, and one student scores exactly 3.0, to maximize the variance.

The mean of the marks will also be:

$$\overline{x}=\frac{1}{n}\cdot\left(\frac{(n-1)}{2}\cdot 1.0+3.0+\frac{(n-1)}{2}\cdot 5.0\right)=\frac{1}{n}\left(\frac{n-1}{2}\cdot 6.0+3.0\right)=3.0\,. \qquad (3\text{-}3)$$

For this mean, the variance is:

$$s^2(\overline{x}=3)=\frac{1}{n}\cdot\left(\frac{(n-1)}{2}\cdot(1.0-3.0)^2+(3.0-3.0)^2+\frac{(n-1)}{2}\cdot(5.0-3.0)^2\right)=$$

$$=\frac{(n-1)\cdot 4}{n}=\frac{4n-4}{n}=4-\frac{4}{n} \qquad (3\text{-}4)$$

and the standard deviation $s=\sqrt{4-\frac{4}{n}}$ so that $\lim\limits_{n\to\infty}s=2$ as for the case of even numbers. As an example: if n=15 students participate in the examination, the maximal standard deviation of the marks would be about s=1.9322, and if n= 55 the maximal s would be 1.9817.

To explore the maximal standard deviation for every mean $1.0\le\overline{x}\le 5.0$, the case of mean 1.0 is considered first; here, all students score $x_i=1.0$ (i=1, …, n). To keep the maximal variance, we now change the mark of only one student by δ ($0\le\delta\le 4$). A change of $\delta=4$ signifies that this student scores 5.0 while the rest of the students score 1.0. Now we introduce the variable z, which counts the number of students with the mark 5.0. At the moment, z=1. From this point on, the mark of one more student who had scored 1.0 will be changed by δ, until (for $\delta=4$) two students have 5.0 and z=2. From this the maximal standard deviation can be described until z=n-1 students score 5.0 and the mark of the last student has been changed by $\delta=4$. This is equivalent to z=n students with 5.0 and $\delta=0$.
Analogously to equations (3-1) and (3-2) respectively, the mean and variance are:

$$\boxed{\overline{x}(n,z,\delta)=\frac{1}{n}\cdot\left((n-z-1)\cdot 1.0+(1.0+\delta)+z\cdot 5.0\right)=1+\frac{4\cdot z+\delta}{n}} \qquad (3\text{-}5)$$

and

$$s^2(n,z,\delta)=\frac{1}{n}\cdot\left((n-z-1)\cdot(1-\overline{x}(n,z,\delta))^2+((1+\delta)-\overline{x}(n,z,\delta))^2+z\cdot(5-\overline{x}(n,z,\delta))^2\right).$$

With the mean $\overline{x}(n,z,\delta)=1+\frac{4\cdot z+\delta}{n}$ and some transformations, the maximal variance:

$$\boxed{s^2(n,z,\delta)=(n-1)\cdot\left(\frac{\delta}{n}\right)^2+\frac{16\cdot(n-z)\cdot z}{n^2}-\frac{8\cdot\delta\cdot z}{n^2}} \qquad (3\text{-}6)$$

can be computed for every n = 1, 2, 3, … , z = 0, …, n-1 and $0\le\delta\le 4$.

In Figures 3-1 and 3-2, the mean of the evaluation marks for a term and the standard deviation are depicted for different numbers n of participants. With the mean 3.0, the differences described for the maximal variance for even and odd numbers n can be seen (see equations (3-2) and (3-4)).

To construct the complete frontier, equations (3-5) and (3-6) are used.

The following table 3-1 applies equations (3-5) and (3-6) to show examples for some n used in Figure 3-2:

n	**z**	**δ**	**mean**	**variance**	**std. dev.**
2	1	0	3.0000	4.0000	2.0000
3	1	2	3.0000	2.6667	1.6330
5	0	0	1.0000	0.0000	0.0000
5	2	2	3.0000	3.2000	1.7889
5	4	4	5.0000	0.0000	0.0000
6	3	0	3.0000	4.0000	2.0000
15	7	2	3.0000	3.7333	1.9322
999	499	2	3.0000	3.9960	1.9990

Table 3-1: Some examples for mean $\overline{x}(n,z,\delta)$ and variance $s^2(n,z,\delta)$

To discuss the average mark in an examination, the reduced maximal standard deviation has to be respected for small numbers of participants ($n \le 15$). For bigger numbers, it is sufficient to use the maximal standard deviation depicted by the black line (n=1000). The line can be computed from equations (3-6) and (3-5). First, equation (3-6) is used with δ=0 to get the frontier of the maximal variance where the only possible marks are 1.0 and 5.0. The result is:

$$s^2(n,z) = \frac{16 \cdot (n-z) \cdot z}{n^2} = 16 \cdot \frac{z}{n} - 16 \cdot \left(\frac{z}{n}\right)^2 . \qquad (3\text{-}7)$$

The variable z moves by discrete steps: z = 1, 2, …, n. The variance s^2 in equation (3-7) depends on $z/n \in [0;1]$. The mean described by equation (3-5) for δ=0 is:

$$\overline{x}(n,z) = 1 + 4 \cdot \frac{z}{n} \qquad (3\text{-}8)$$

or, rewritten for z/n:

$$\frac{z}{n} = \frac{\overline{x}-1}{4} . \qquad (3\text{-}9)$$

From equation (3-9) the variance of equation (3-7) depends on z/n and can be transformed into:

$$s^2(\overline{x}) = 16 \cdot \frac{\overline{x}-1}{4} - 16 \cdot \left(\frac{\overline{x}-1}{4}\right)^2 = 4 \cdot \overline{x} - 4 - \overline{x}^2 + 2 \cdot \overline{x} - 1 =$$

$$\boxed{s^2(\overline{x}) = -\overline{x}^2 + 6 \cdot \overline{x} - 5}, \text{ with } 1.0 \le \overline{x} \le 5.0 . \qquad (3\text{-}10a)$$

In the final chapter below, a more general method will be used to get this frontier function. The examples in Table 3-2 show some applications of equation (3-10a). As in Figures 3-1 and 3-2, when the mean is 1.0 or 5.0 the standard deviation s = 0.00. The mean of 3.0 corresponds to the maximal standard deviation s=2.00. Above and below this point, the standard deviation shrinks.

To estimate an "inner" curve, which describes for small n the maximal standard deviation (or, better, the minima of the maximal standard deviation) the equation (3-6) can again be used. To get the middle point of the different curve segments, the parameter δ is fixed at δ = 2. Furthermore, equation (3-5) has to be rewritten as $z = (n \cdot (\overline{x}-1) - \delta)/4$ (so $z = (n \cdot (\overline{x}-1) - 2)/4$) and applied to

mean	variance	std. dev.
5.00	0.00	0.00
4.50	1.75	1.32
4.00	3.00	1.73
3.50	3,75	1.94
3.00	4.00	2.00
2.50	3.75	1.94
2.00	3.00	1.73
1.30	1.11	1.05
1.00	0.00	0.00

Table 3-2: The maximal mean-standard-deviation frontier using equation (3-10a)

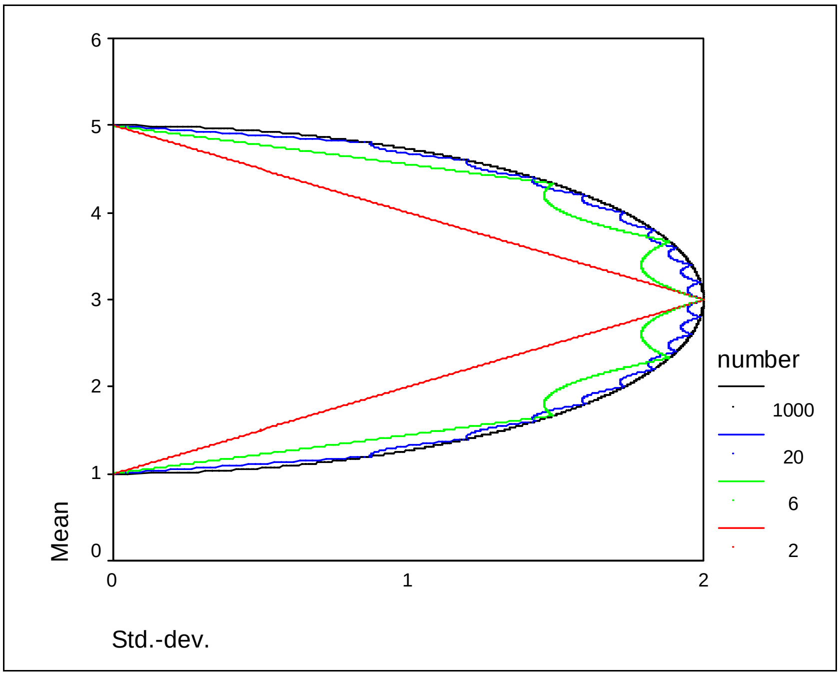

Figure 3-1: Maximal standard deviation in the case of the even numbers n= 2, 6, 20, 1000

substitute z in equation (3-6). After some transformations, the maximal variance for small n – characterized by this “inner” curve – is:

$$\boxed{s_I^{\,2}(\bar{x}) = -\bar{x}^2 + 6\cdot\bar{x} - 5 - \frac{4}{n}}\text{, with } (n \geq 2) \text{ and } 1.0 \leq \bar{x} \leq 5.0\,. \qquad \text{(3-10b)}$$

Equation (3-10b) is an approximation, which shifts the variance in equation (3-10a) by 4/n, to give an idea of where the maximal variance lies in the case of small n. For these cases, the variance lies in the area between the curves (3-10a) and (3-10b). For extreme examinations, with either only n scores of 1.0 or only n scores of 5.0, the approximation of (3-10b) is not applicable , especially for very small n. This is because of the use of $\delta = 2$ in the construction of the “inner” curve. For a mean of 1.0 or 5.0, s_I^2 would be negative. In Table 3-3, the differences in the standard deviation for $n=\infty$ and very small n can be seen. With a mean 3.0 for $n=\infty$ the standard deviation $s_I= 2.00$, but this measure for n=5 or n=15 is reduced to $s_I=1.79$ (or, respectively, $s_I=1.93$). With a mean of 1.30, these differences are greater. For the case of $n=\infty$ equation (3-10b) is identical to equation (3-10a). The inner curves for n=5, n=15 and $n=\infty$ can be seen in Figure 3-3. The dotted line signifies the mean of 1.30 in the chart.

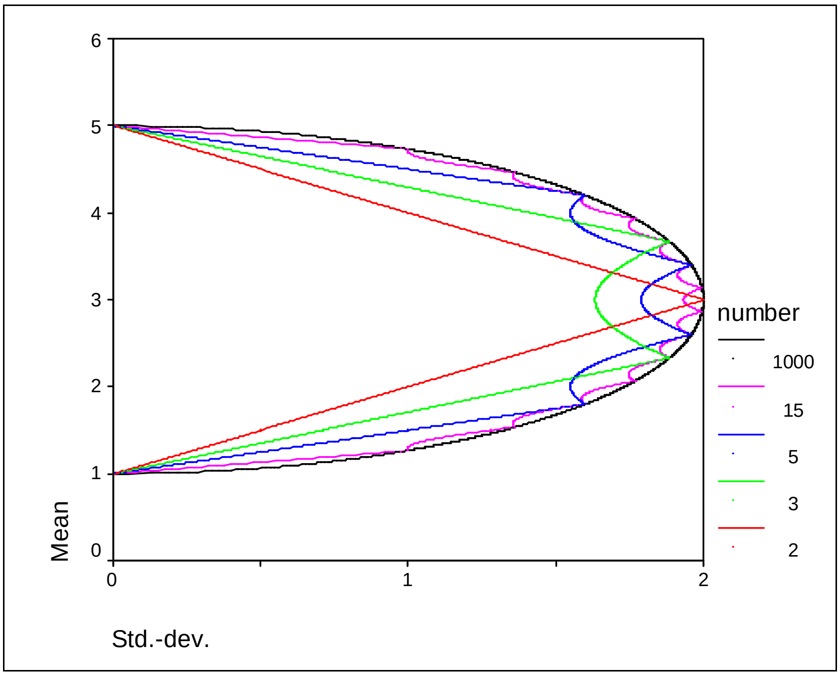

Figure 3-2: Maximal standard deviation in the case of the odd numbers n= 3, 5, 15

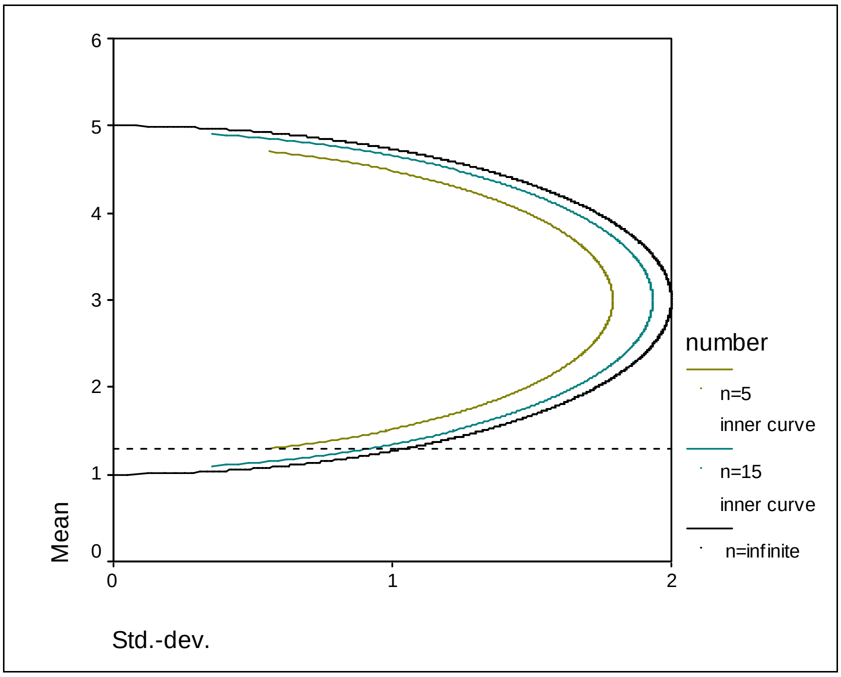

Figure 3-3: Inner curves of the maximal standard deviation for n=5 and n=15

number n	mean	variance	std. dev.
∞	**3.00**	**4.00**	**2.00**
100	3.00	3.96	1.99
50	3.00	3.92	1.98
20	3.00	3.80	1.95
15	3.00	3.73	1.93
5	3.00	3.20	1.79
∞	**1.30**	**1.11**	**1.05**
100	1.30	1.07	1.03
50	1.30	1.03	1.01
20	1.30	0.91	0.95
15	1.30	0.84	0.92
5	1.30	0.31	0.56

Table 3-3: The maximal mean-standard deviation frontier using equation (3-10b) (for mean 1.3 see dotted line in Figure 3-3)

4. The maximal standard deviation with marks from 1.0 to 6.0

Some schools use marks from x_i=1.0 to 6.0 for the i=1, …, n participants. To explore the maximal standard deviation for every mean $1.0 \le \overline{x} \le 6.0$ in this case, the same procedure as that of chapter 3 is used. The parameter δ that describes the change in one student's mark lies in the interval $\delta \in [0;5]$. For δ=5, one mark changes from 1.0 to 6.0. The variable z counts the number of students with a mark of 6.0, as in chapter 3. The equation (3-5) for the mean is modified to:

$$\boxed{\overline{x}(n,z,\delta) = 1 + \frac{5 \cdot z + \delta}{n}} \qquad (4\text{-}1)$$

and the maximal variance for this mean can be derived, in the same way as in chapter 3, as:

$$\boxed{s^2(n,z,\delta) = (n-1) \cdot \left(\frac{\delta}{n}\right)^2 + \frac{25 \cdot (n-z) \cdot z}{n^2} - \frac{10 \cdot \delta \cdot z}{n^2}}, \qquad (4\text{-}2)$$

with n = 1, 2, 3, … , z = 0, …, n-1 and $0 \le \delta \le 5$.

For the mean of 3.5, the maximal variance s^2 is 6.25 and the standard deviation s=2.5 (e.g. n=6, z=3, δ=0). The maximal variance for odd numbers n converges to this value as n increases (e.g.: s^2=5.0 and s=2.23607 for n=5, z=2, δ=2.5). In Table 4-1 the equations (4-1) and (4-2) are used to give some examples.

n	**z**	**δ**	**mean**	**variance**	**std. dev.**
2	1	0	3.50	6.2500	2.5000
3	1	2.5	3.50	4.1667	2.0412
5	0	0	1.00	0.0000	0.0000
5	2	2.5	3.50	5.0000	2.2361
5	4	5	6.00	0.0000	0.0000
6	3	0	3.50	6.2500	2.5000
15	7	2.5	3.50	5.8333	2.4152
999	499	2.5	3.50	6.2437	2.4987

Table 4-1: Some examples for mean $\overline{x}(n,z,\delta)$ and variance $s^2(n,z,\delta)$

In the same way as in chapter 3, equations (4-1) and (4-2) can be transformed (with δ=0) to get the maximal variance s^2 as a function of the mean:

$$\boxed{s^2(\bar{x}) = -\bar{x}^2 + 7\cdot\bar{x} - 6}, \text{ with } 1.0 \le \bar{x} \le 6.0 \qquad (4\text{-}3a)$$

and the maximal variance for small n described by the “inner” curve (with δ=2.5 and $z = (n\cdot(\bar{x}-1) - 2.5)/5$) is

$$\boxed{s^2(\bar{x}) = -\bar{x}^2 + 7\cdot\bar{x} - 6 - \frac{6.25}{n}}, \text{ with } 1.0 \le \bar{x} \le 6.0\,. \qquad (4\text{-}3b)$$

mean	variance	std.dev.
6.00	0.00	0.00
5.55	2.25	1.50
5.00	4.00	2.00
4.50	5.25	2.29
4.00	6.00	2.45
3.50	6.25	2.50
3.00	6.00	2.45
2.50	5.25	2.29
2.00	4.00	2.00
1.50	2.25	1.50
1.30	1.41	1.19
1.00	0.00	0.00

number n	mean	variance	std.dev.
∞	**3.50**	**6.25**	**2.50**
100	3.50	6.19	2.49
50	3.50	6.13	2.47
20	3.50	5.94	2.44
15	3.50	5.83	2.42
5	3.50	5.00	2.24
∞	**1.30**	**1.41**	**1.19**
100	1.30	1.35	1.16
50	1.30	1.29	1.13
20	1.30	1.10	1.05
15	1.30	0.99	1.00
5	1.30	0.16	0.40

Table 4-2: Some examples for equations (4-3a) (left side) and (4-3b) (right side)

5. Target-Shortfall-Probability and skewed distributions

The Target-Shortfall-Probability (TSP) is well known for the case in which the marks x_i are normally distributed (or X ~ N(0,1)). In applications of the TSP for the average marks for an examination, the target τ could be, for example, τ=2.5. At least this mark is needed to get a good examination result. In the case of X ~ N(0,1), the probability that at most α=70% of the students get a good or very good result is:

$$P(X<\tau) \le \alpha \text{ so } P(X< 2.5) \le 0.70. \qquad (5\text{-}1a)$$

The target τ of the inequality (5-1a) can be normalized by the mean μ and standardized by the standard deviation σ of the variable X by

$(\tau - \mu) / \sigma \le z_\alpha$ with z_α the abscissa value of the N(0,1) probability distribution. (5-1b)

The inequality (5-1b) can be rearranged to give the expected value of X

$$\mu \ge \tau - z_\alpha\cdot\sigma. \qquad (5\text{-}1c)$$

In the inequality (5-1c) the target τ=2.5 and $z_{0.7}$ = 0.5244 gives a borderline for the mean and standard deviations. All μ-σ-combinations above this line fulfil the condition that at most α=70% of the students get a good or very good result. In this example, the line is:

$$\mu = 2.5 - 0.5244\cdot\sigma. \qquad (5\text{-}1d)$$

This line can be integrated into the graphs from chapters 3 and 4. Every examination that has a μ-σ-combination below this line will have more than 70% of good and very good marks.

The restriction of the marks to the interval [1;5] means that in many cases the normal distribution will be deformed to a skewed distribution. Therefore, the line (5-1d) has to be adapted. An alternative line can be found using the extreme cases of the curve of the graphs of chapter 3. If 70% score 1.0 and

30% score 5.0, the examination will have a mean of $\bar{x} = 2.2$ and a standard deviation of s=1.8330. For this standard deviation and mean, exactly 70% get a good or very good result. With no standard deviation (s=0), the mean has to be 2.5. From these two points, the gradient of the line can be computed. The line starts at the point (μ;σ) = (2.5, 0) and ends at the point (2.2, 1.8330). The computation is:

$$(2.2 – 2.5) / (1.8330 – 0) = -0.1637.$$

This procedure is applied to the other probabilities and the targets 2.5 and 3.0, in Table 5-1. Compared with the gradient $-z_\alpha$ for the N(0,1) distribution (see inequalities (5-1b) to (5-1d)), the gradient is reduced. This is also shown in Figure 5-1, in which the target τ=2.5 was used for the TSP lines.

probability of being below the target	mean $\bar{x}$	std. dev. s	gradient for target τ=2.5	gradient for target τ=3.0
90	1.4	1.200000	-0.9167	-1.3333
80	1.8	1.600000	-0.4375	-0.7500
70	2.2	1.833030	-0.1637	-0.4364
60	2.6	1.959592	0.0510	-0.2041
50	3.0	2.000000	0.2500	-0.0000
40	3.4	1.959592	0.4593	0.2041
30	3.8	1.833030	0.7092	0.4364
20	4.2	1.600000	1.0625	0.7500
10	4.6	1.200000	1.7500	1.3333

Table 5-1: Estimated gradients of the TSP line for skewed distributions

The “Skew-TSP” lines in Figure 5-1 are correct only at the beginnings and ends of the lines where a one-point distribution exists. These “Skew-TSP” lines may be nonlinear. They signify that the TSP line will move upwards and be deformed when the normal distribution is restricted at the lower side of the scale.

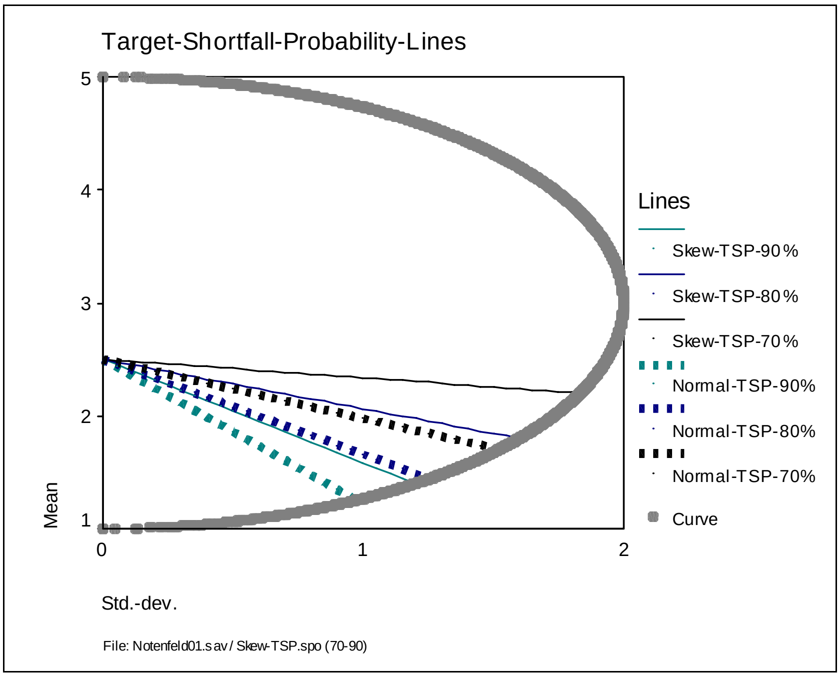

Figure 5-1: TSP lines for 70%, 80% and 90% and “Skew-TSP” lines that coincide with the TSP in extreme situations (as they do in the curve).

6. Minimal standard deviation when marks are discrete numbers

Sometimes the marks are additionally restricted to discrete numbers like 1.0, 1.3, 1.7, 2.0, 2.3 etc. This constraint means that a grade point average of 1.8 cannot happen with zero standard deviation. The minimal dispersion can only occur when the average mark is one of the discrete numbers from the list above. To find the standard deviation curve between two discrete zero dispersion points a and b (a≠b) (e.g. a=1.3, b=1.7) that are direct neighbours, the variance function s^2 is used. For this, the function of the average mark is expressed as a function of a variable y that is the proportion of students with the smaller mark a (so that (1-y) is the proportion with the higher mark):

$$\bar{x} = y \cdot a + (1 - y) \cdot b \, . \qquad (6\text{-}1)$$

The mean equation (6-1) can be rewritten with the proportion y as the subject:

$$y = \frac{\bar{x} - b}{a - b} \, . \qquad (6\text{-}2)$$

The variance, dependent on the mean, is:

$$s^2(\bar{x}) = y \cdot [a - \bar{x}]^2 + (1 - y) \cdot [b - \bar{x}]^2 \qquad (6\text{-}3)$$

With the equations (6-1) and (6-2) the variance can be expressed as a function of the mean and the marks a and b:

$$s^2(\bar{x}) = \frac{\bar{x} - b}{a - b} \cdot [a - \bar{x}]^2 + \left(1 - \frac{\bar{x} - b}{a - b}\right) \cdot [b - \bar{x}]^2 \qquad (6\text{-}4)$$

Equation (6-4) can be transformed[8] into equation (6-5):

$$s^2(\bar{x}) = -\bar{x}^2 + (a + b) \cdot \bar{x} - ab \, . \qquad (6\text{-}5)$$

The variance function (6-5) is a generalisation of the special case of equation (3-10a) where a=1 and b=5 or of equation (4-3a) where a=1 and b=6.

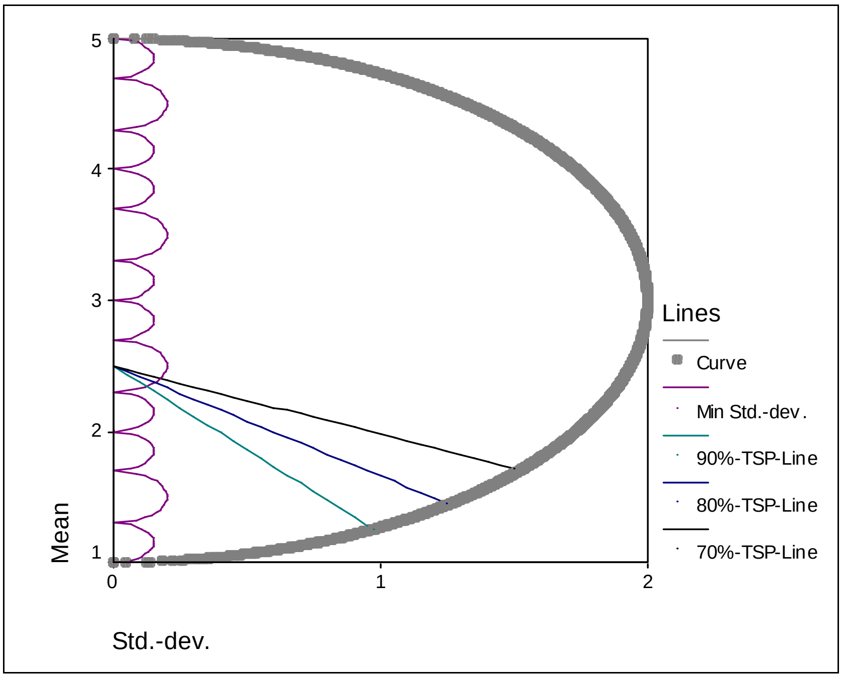

Figure 6-1: Minimal standard deviation when marks are discrete, with TSP lines for 70%, 80% and 90%.

[8] The transformation is explained in the appendix.

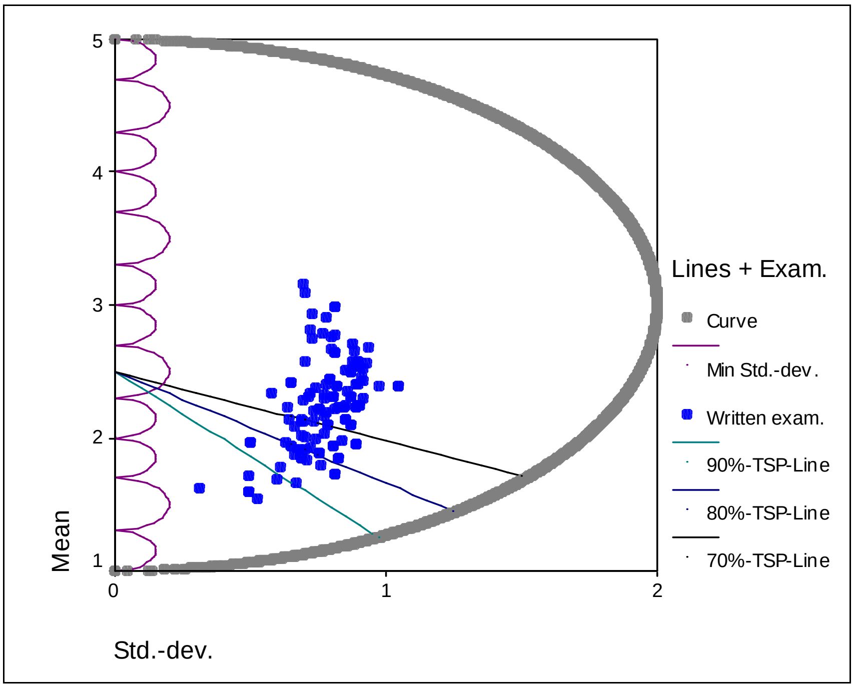

Figure 6-2: Mean and standard deviation of average marks of some written examinations in the mean-standard deviation space.

For examinations with small numbers of participants, only some points of the variance function can be realized. When, for example, only two students participate in an examination, three possibilities may occur: both will get a or b, with $s^2=0$, or one will get a and the other b, with $s^2> 0$ computed from equation (6-5). Between the marks a and b must not lie another mark c with a<c<b. Then the variance of the marks a and b will be minimal for the interval [a, b]. This means that equation (6-5) gives the minimal variance for two neighbours a and b.

Figure 6-1 shows the left limitation of the mean-standard deviation space when the marks are not continuous. In the example, the marks have the following values: 1.0, 1.3, 1.7, 2.0, 2.3, 2.7, 3.0, 3.3, 3.7, 4.0, 4.3, 4.7, 5.0. Furthermore, some TSP lines are integrated.

In Figure 6-2 the standard deviation and the average marks for some written examinations are depicted. Only five examinations are below the 90%-TSP line. The majority lie above this line. There are two examinations with an average mark worse than 3.0. The mean-standard deviation framework is useable when the relative position of a few examinations needs to be shown. Besides the visibility of the information about the level of the average marks and their position relative to the TSP line, high and low standard deviations can also be recognized. Low standard deviation signifies that only some marks were used in the evaluation of the examination. Near the violet curves on the left-hand side, the average mark was built from only one or two different marks. On the other side, high standard deviation results from a distribution of marks with two peaks. In the extreme position of the curve, only the marks 1.0 and 5.0 were given.

7. Conclusion and suggestions for further research

The instrument that is developed in this paper is not an analytical one. It is more an instrument for visualisation that shows the possible area in which the mean and standard deviation of a restricted variable can lie. The example in this paper is concerned with the average mark for an examination and the dispersion of the marks. Neither a high nor a low dispersion of marks should be an aim when marking examinations, and the same applies to a high or low grade point average. These extreme positions should occur when they are reasonable.
The impact of the skewness of the distribution on the Target-Shortfall-Probability needs some further investigation. In this paper, only an upper bound for the TSP line was proposed.

8. References

[1] Bamberg, G., Baur, F., Krapp M., (2012): Statistik, Oldenbourg.
[2] Bouza Herrera, C., Schubert, L. (2003): The estimation of Biodiversity and the Characterization of the Dynamics: An Application to the Study of Pest, Rev. Mat. Estat., Sao Paulo, vol. 21, pp. 85-98.
[3] Grundmann, W., Luderer, B., (2003): Formelsammlung Finanzmathematik, Versicherungsmathematik, Wertpapieranalyse, Teubner, Stuttgart.
[4] Schubert, L., (2002): Portfolio Optimization with Target-Shortfall-Probability-Vector, Economics Analysis Working Papers, Vol. 1 No 3, 2002, ISSN15791475.
[5] Wallace, D. F., (2009): Infinity Jest, German edition, rororo, 2009.

Appendix

Equation (6-4) is

$$s^2(\bar{x}) = \frac{\bar{x}-b}{a-b}\cdot[a-\bar{x}]^2 + \left(1-\frac{\bar{x}-b}{a-b}\right)\cdot[b-\bar{x}]^2 .$$

Rewriting this equation has the result:

$$s^2(\bar{x}) = \frac{\bar{x}a^2 - 2a\bar{x}^2 + \bar{x}^3 - 2ba^2 + 2ab\bar{x} - b\bar{x}^2 - \bar{x}b^2 + 2b\bar{x}^2 - \bar{x}^3 + b^3 - 2b^2\bar{x} + b\bar{x}^2}{(a-b)} + b^2 - 2b\bar{x} + \bar{x}^2$$

which can be reduced by the compensating elements $+\bar{x}^3 - b\bar{x}^2 - \bar{x}^3 + b\bar{x}^2$ in the left fraction. The remaining term is ordered by $\bar{x}^2$, $\bar{x}$ and the constant element to

$$s^2(\bar{x}) = \bar{x}^2\cdot\left(\frac{-2a+2b}{a-b}+1\right) + \bar{x}\cdot\left(\frac{a^2+2ab-b^2-2b^2}{a-b}-2b\right) + \frac{-ba^2+b^3}{a-b} + b^2 .$$

With +1 in the left summand substituted by (a-b)/(a-b) and analogously with b^2 by $b^2\cdot$(a-b)/(a-b) in the right summand, the equation can be simplified to:

$$s^2(\bar{x}) = -\bar{x}^2 + \bar{x}\cdot\left(\frac{a^2+2ab-3b^2-2ba+2b^2}{a-b}\right) + \frac{-ba^2+b^3+ab^2-b^3}{a-b} .$$

Uniting some summands in both fractions gives the function (6-5) for the maximal variance:

$$s^2(\bar{x}) = -\bar{x}^2 + \frac{a^2-b^2}{a-b}\cdot\bar{x} + \frac{ab^2-a^2b}{a-b} .$$

The factor $\frac{a^2-b^2}{a-b} = \frac{(a+b)\cdot(a-b)}{a-b} = (a+b)$ and $\frac{ab^2-a^2b}{a-b} = \frac{-ab\cdot(a-b)}{a-b} = -ab$ substituted give

the function (6-5) for the maximal variance:

$$s^2(\bar{x}) = -\bar{x}^2 + (a+b)\cdot\bar{x} - ab .$$