Multivariate Inverted Normal Loss Functions

David Drain and Elizabeth Cudney

University of Missouri-Rolla

Missouri University of Science and Technology

...a micron is as good as a mile

Loss functions

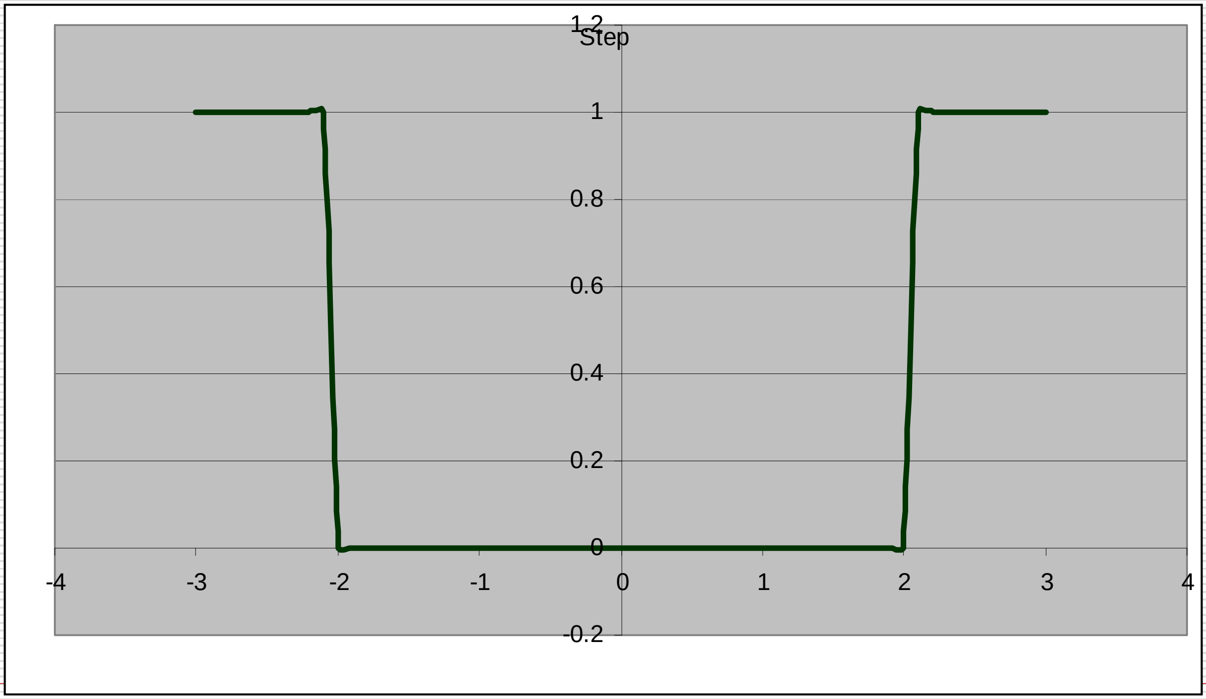
- Quantify the *loss* (cost) of variation from target

Specification Limits

- Hard limits determined by customer:
  - Material inside the limits is “good”
  - Material outside the limits is “bad”

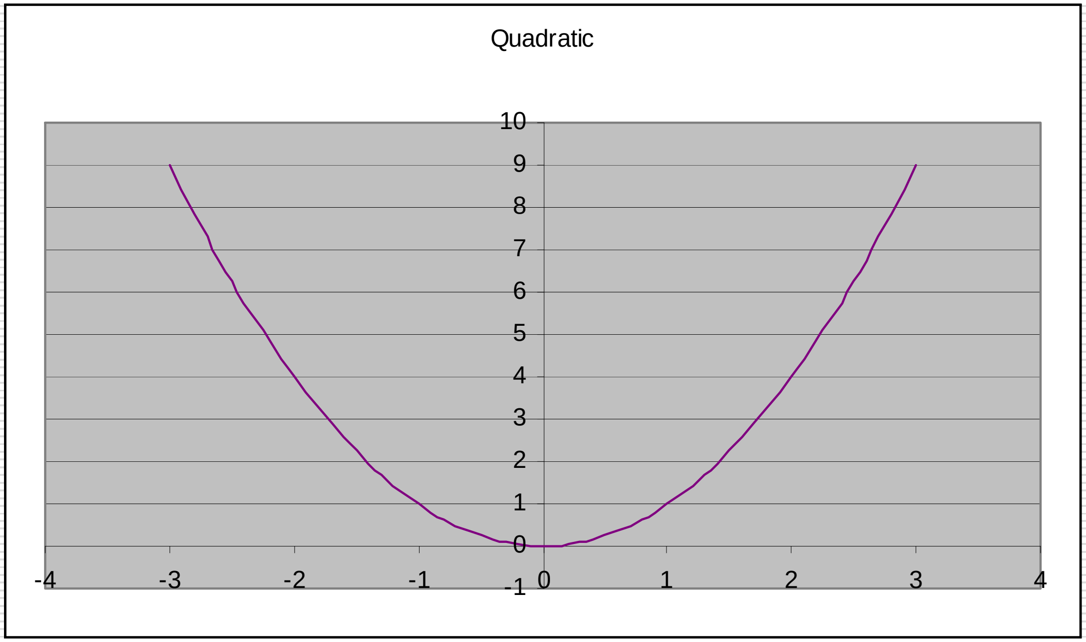
Step function loss

- The use of specification limits for product screening implies *step function loss*:



Quadratic Loss

- Taguchi (and others) recognized that step function loss was unrealistic, and proposed quadratic loss:

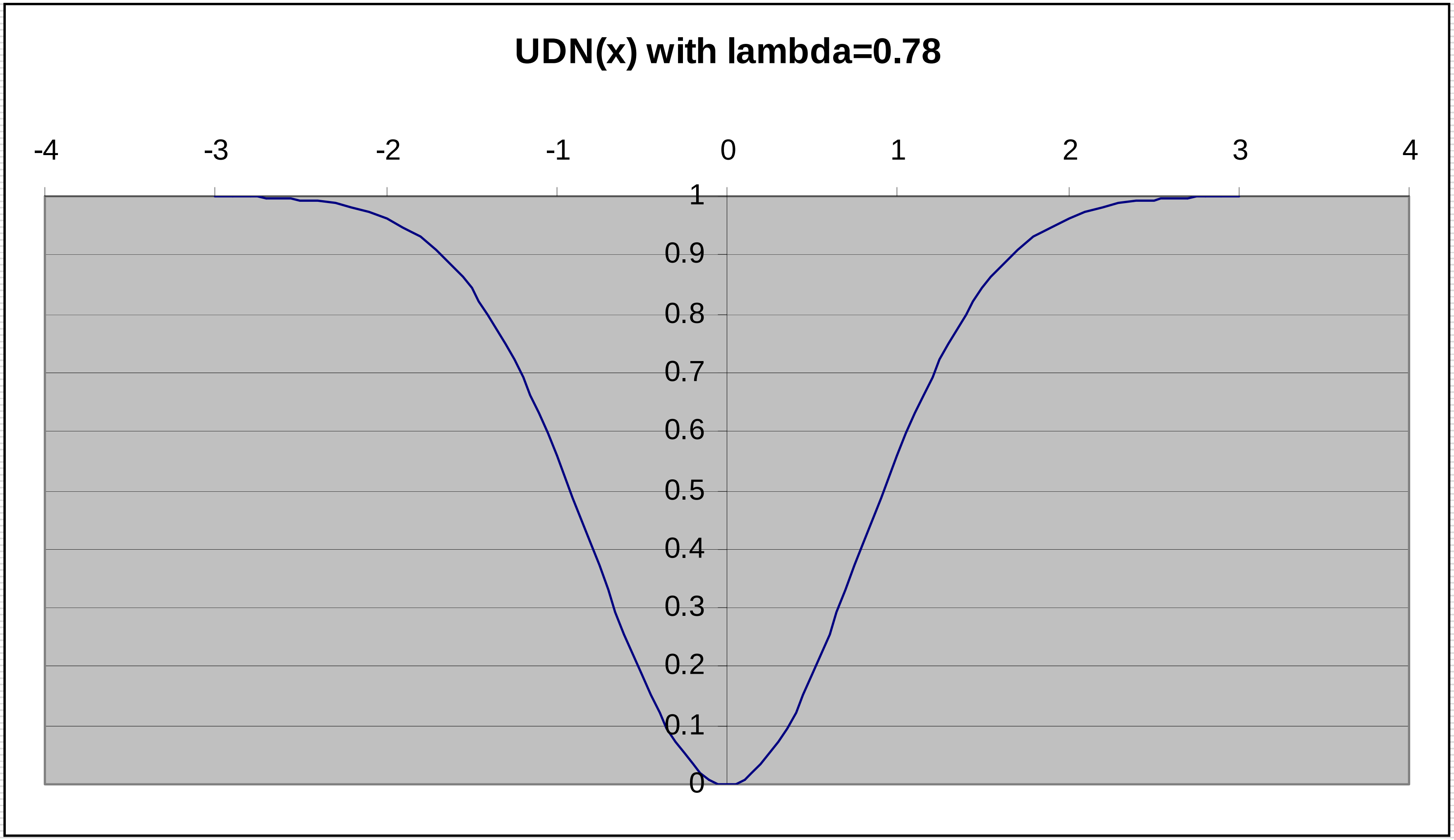


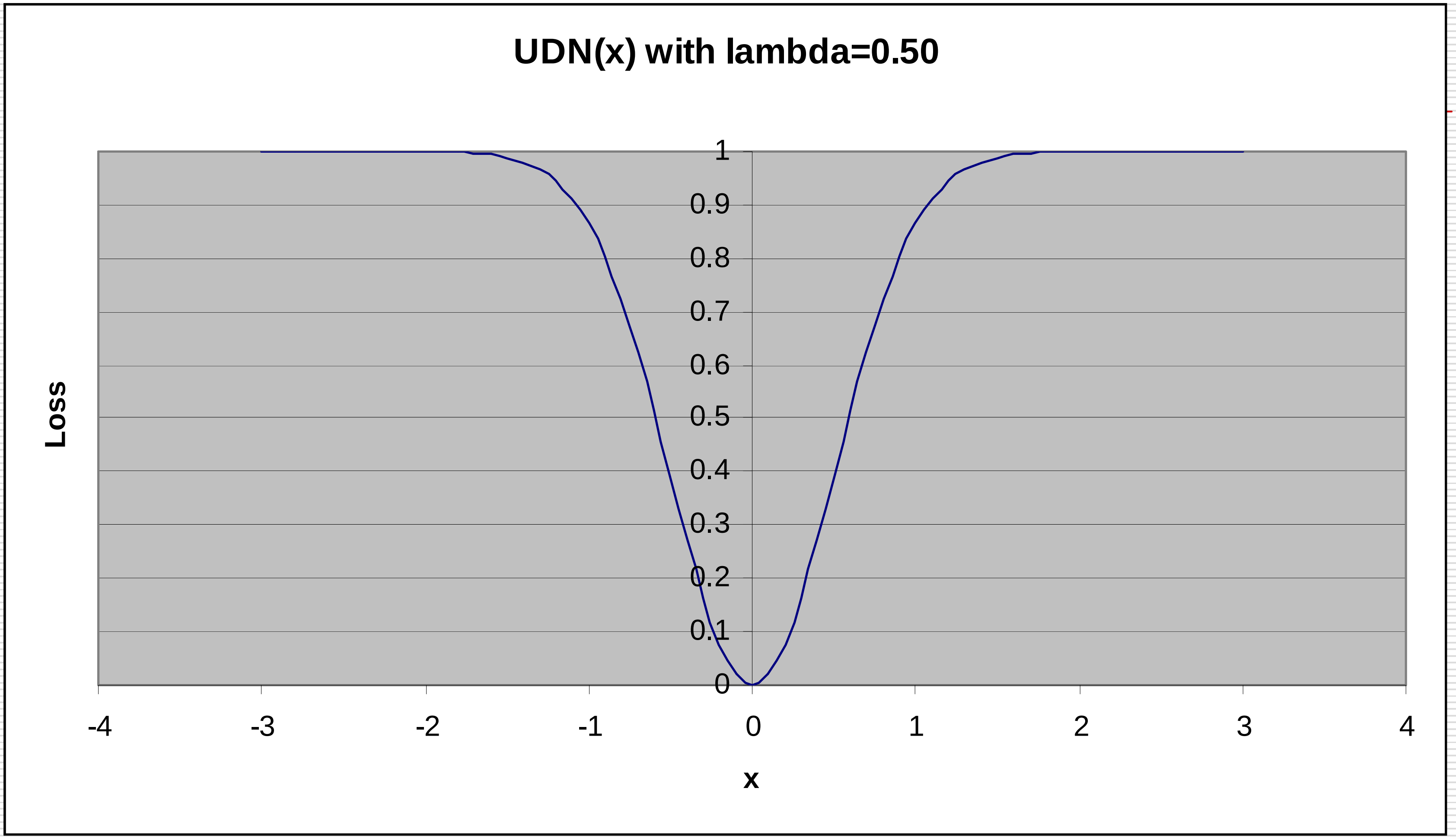
Inverted Normal Loss

- Bounds the loss between zero and one
- Recognizes that fact that all material too far from target is equally bad
- Has very useful mathematical properties:
  - Bounded and infinitely differentiable
  - Scalable to model real losses
  - Simple closed-form solution for expected value with normally distributed processes
  - Extends to multivariate case with ease

$$L(x;\tau,\lambda)=1-e^{-\frac{(x-\tau)^2}{2\lambda^2}}$$

- INL is a scaled inverted probability density function for the normal distribution
- τ is process target, λ is a scale parameter
- Loss is zero at target
- Larger λ gives a less sensitive loss function


UDN(x) with lambda=0.78
-4
-3
-2
-1
0
1
2
3
4
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0



UDN(x) with lambda=0.50
Loss
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-4
-3
-2
-1
0
1
2
3
4
x

Why INLF?

- In the semiconductor industry, a micron is as good as a mile
- Unlike many low-tech industries, losses far from target *can* occur with non-zero probabilty
- A lithography critical dimension 20 nm off-target is no worse than one 5 nm off-target, so they should be assigned the same loss
- Quadratic loss distorts loss computations and leads to sub-optimal decisions

Extensions to simple INLF

- Asymmetric INLFs have similar properties
- Any form of distribution can be used to model the process, but numerical integration may be required for expected values
- INLF properties extend easily to MINLF, where correlations between process variables can have interesting consequences

More general inverted probability loss functions

- Leung and Spiring introduced and developed an entirely new family of loss functions based in inverted pdfs:
  - Beta
  - Gamma
  - Some results for general IPLFs

Parameter estimation

- Best to estimate λ from actual loss data using non-linear regression based on historical data
- If actual loss data is unavailable, or consistency with step-function loss is required, choose λ to give 50% loss at a specification limit:
  - $\lambda = 0.425(USL-LSL)$

Expected Loss

- If the process is normally distributed with mean μ and standard deviation σ, then the average loss from that process will be:

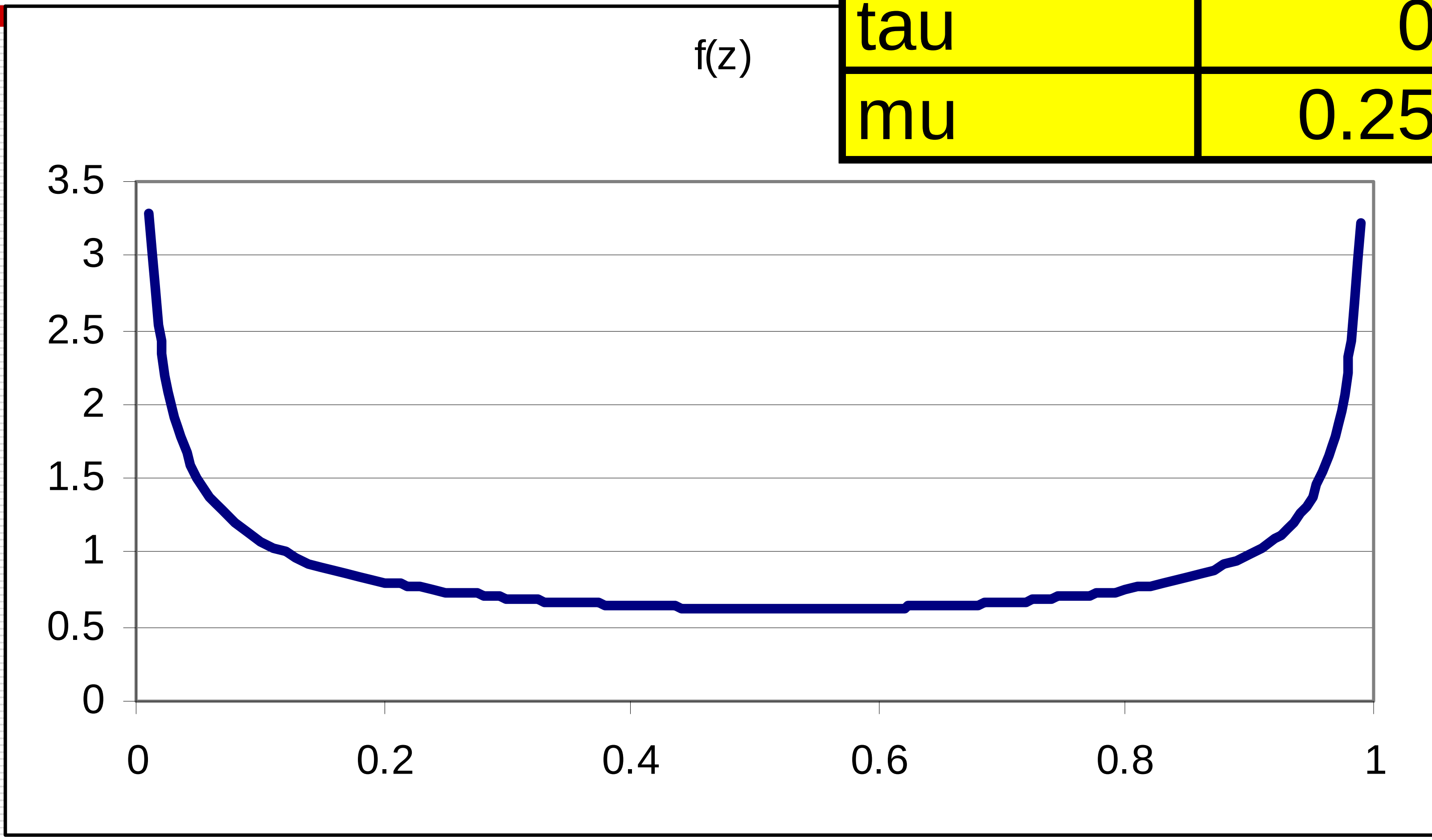
$$1-\frac{\lambda}{\sqrt{\sigma^2+\lambda^2}}e^{-\frac{(\mu-\tau)^2}{2\left(\sigma^2+\lambda^2\right)}}$$

PDF of INLF

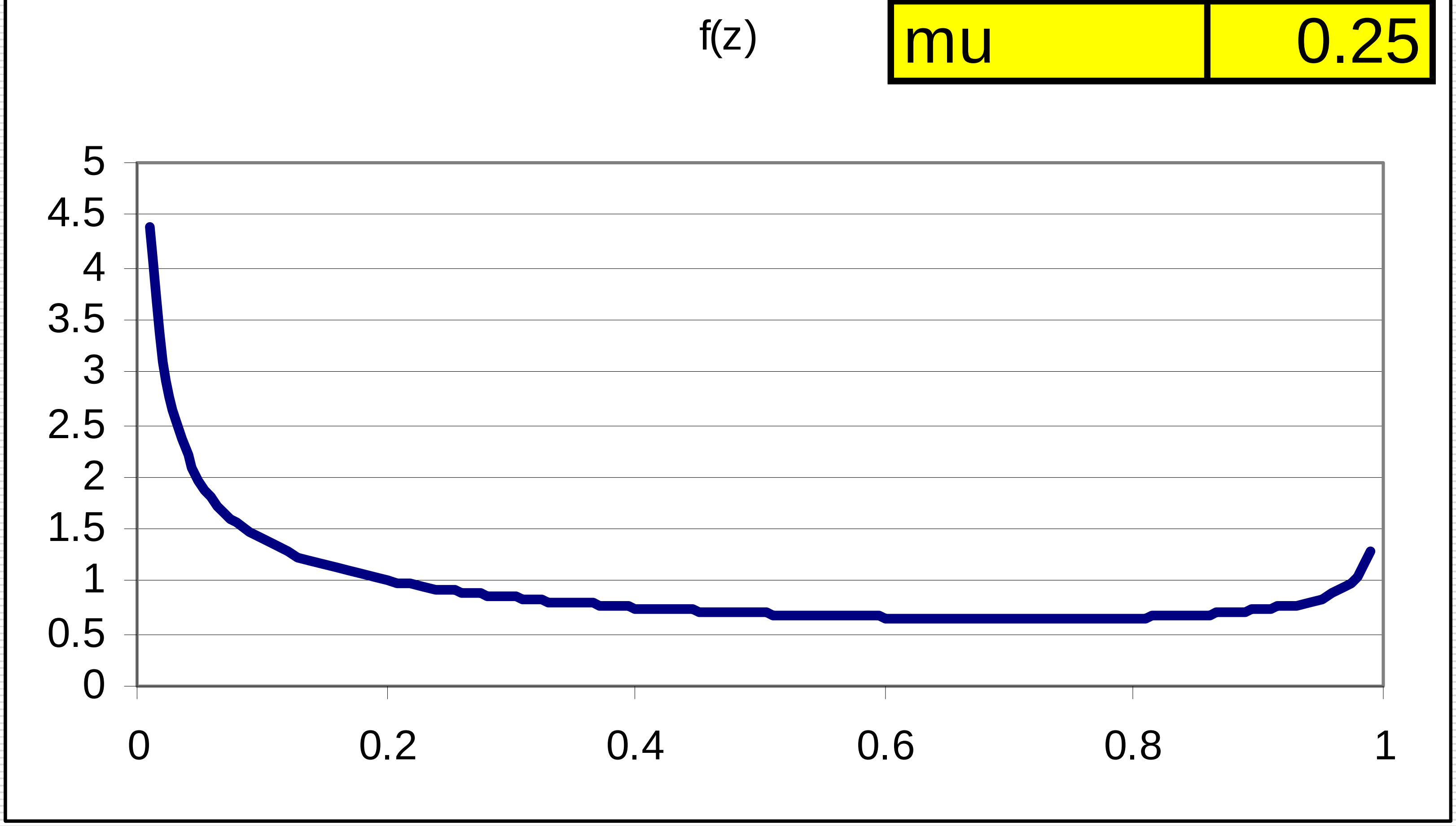
- Through standard transformation methods the loss pdf (with normal processes) is given by:

$$f(z) = \frac{\lambda\left(-\ln(1-z)\right)^{-\frac{1}{2}}}{2(1-z)\sqrt{\pi}\sigma}\left\{ e^{-\frac{\left(\tau+\lambda\sqrt{-2\ln(1-z)}-\mu\right)^2}{2\sigma^2}} + e^{-\frac{\left(\tau-\lambda\sqrt{-2\ln(1-z)}-\mu\right)^2}{2\sigma^2}} \right\}$$

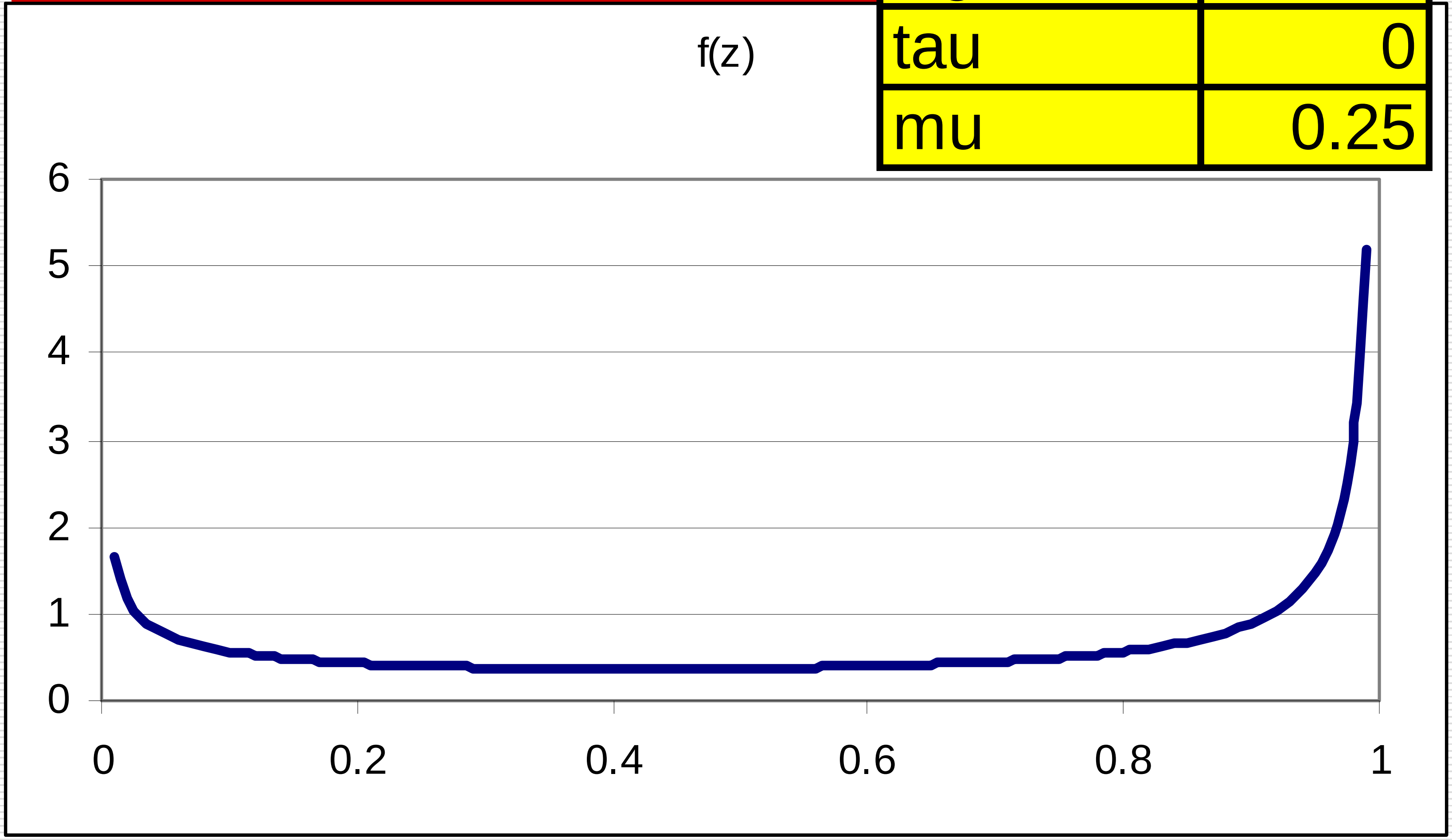
lambda	0.6
sigma	1
tau	0
mu	0.25



lambda	0.8
sigma	1
tau	0
mu	0.25



lambda	0.3
sigma	1
tau	0
mu	0.25



MGF of INLF

- Should be an easy integration
- …or an infinite series based on moments (see Leung and Spiring)
- As far as we know, this is an unsolved problem
- I’ll buy a marguerita for the first person solving this problem

Applications of INLF

- Quantify process quality
- Evaluate alternative process targets
- Evaluate process and equipment changes
- Alternative to squared-error loss for robust regression
- Alternative to squared-error loss in Bayesian estimation

Expected loss and Cpk

- Expected loss gives a truer picture of process health than Cpk
- Deviation from target is always reflected in expected loss
- Atypical process distributions or loss relationships are also comprehended

$$C_{pk} = \min\left\{\frac{USL-\mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right\}$$

$where$

μ = process mean

σ = process standard deviation

USL = upper specification limit

LSL = lower specification limit

C_{pk} and Percent OOS

- For normally distributed process data with two-sided specification limits, there is a simple correspondence between C_{pk} and the percent of material out of specification

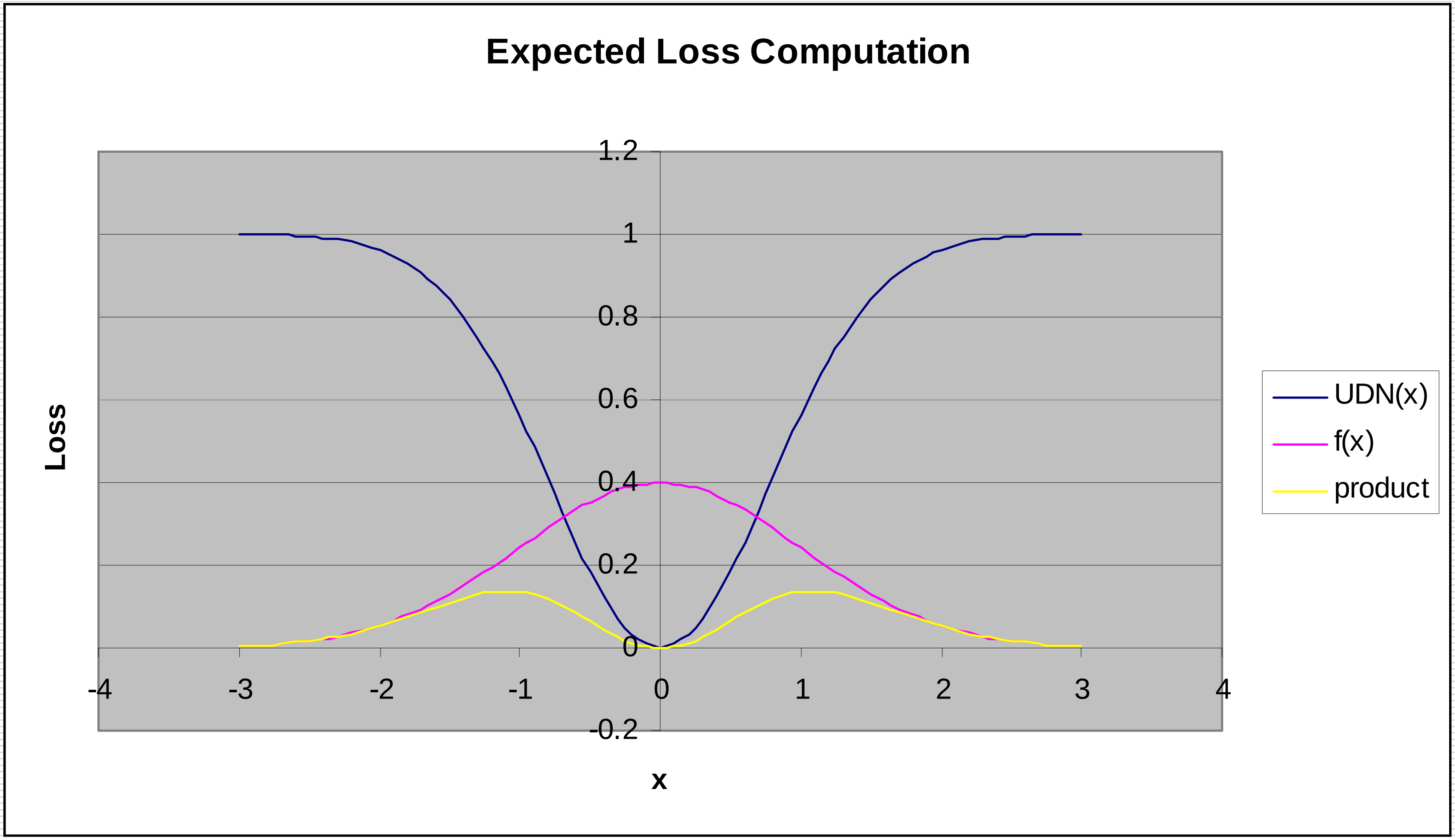
Cpk	Proportion OOS
0.33	3.173E-01
0.67	4.550E-02
1.00	2.700E-03
1.33	6.337E-05
1.67	5.742E-07
2.00	1.980E-09
2.33	2.576E-12
2.67	1.332E-15

C_{pk} is often misused

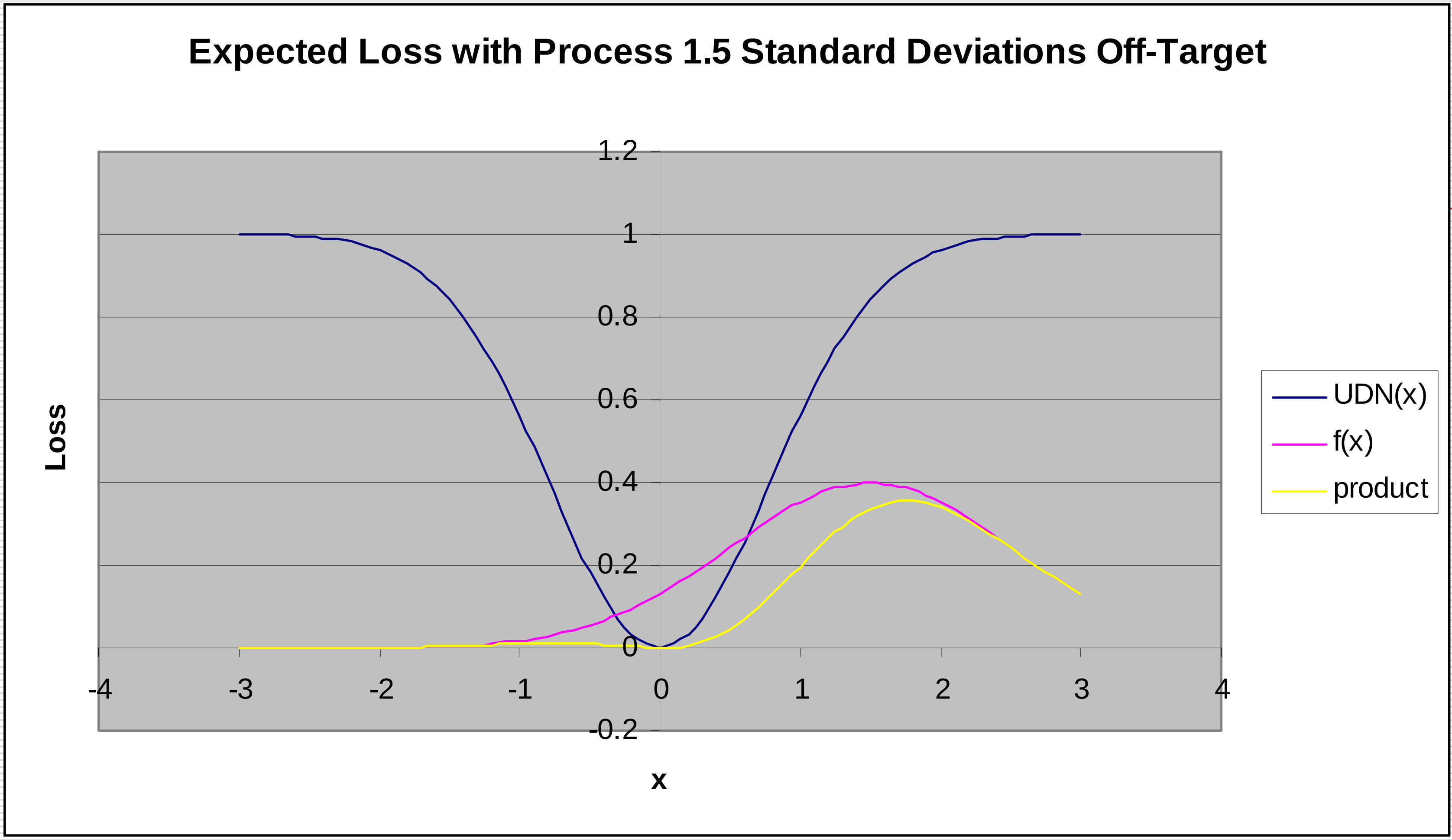
- C_{pk} is applied to non-normally distributed data - probability interpretations no longer apply here
- C_{pk} is applied to processes with one-sided specification limits
- A process can run far off target, but with small variance, and still have acceptable C_{pk}

Comparison of C_{pk} and EL_{INL}

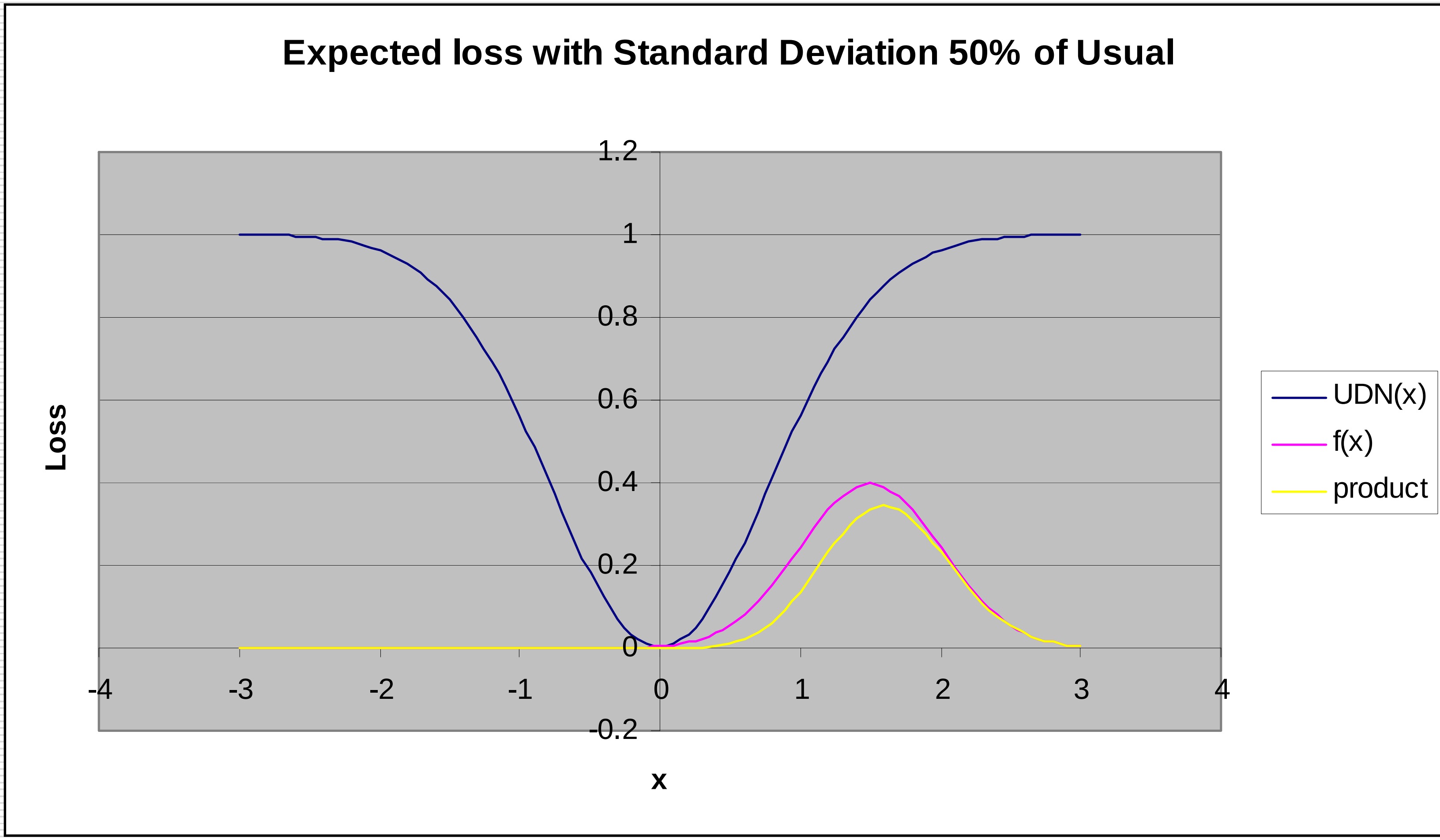
- The expected loss "punishes" deviation from target, even if the process standard deviation is small
- Expected loss can also be computed for other underlying distributions, so is not dependent on the assumption of process normality



Specification limits are at +/- 3



$$EL=0.693, \quad C_{pk}=0.50$$



$$EL=0.773, \ C_{pk}=1.00$$

Multivariate INLF

- First proposed by Drain and Gough, 1996
- Accounts for synergy or antagonism among the process (or noise) variables
- Expected loss (with multivariate normal process) has a simple solution

MINLF definition

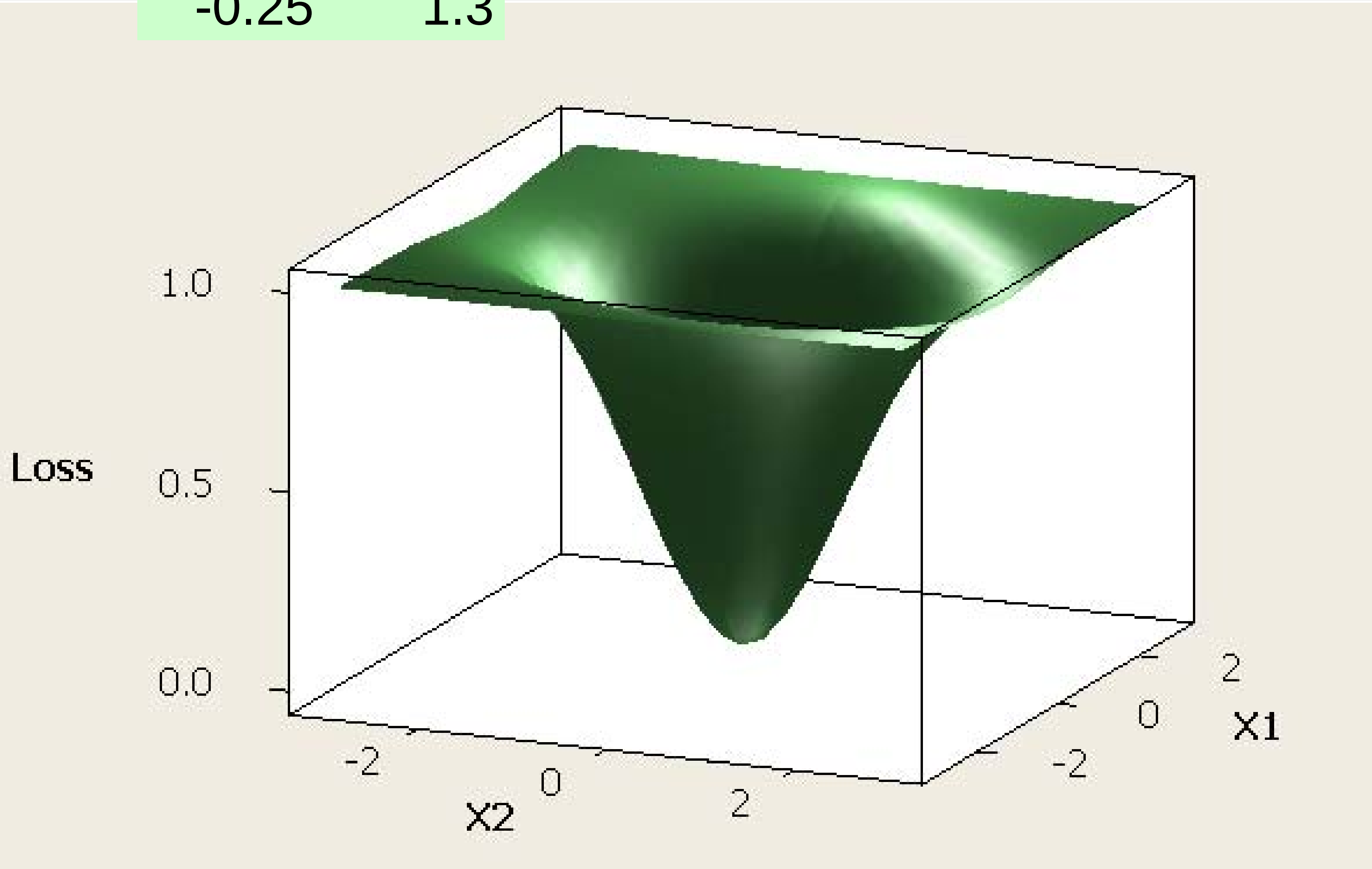
$$L(x;\tau,L) = 1 - e^{-\frac{1}{2}(x-\tau)^T L^{-1}(x-\tau)}$$

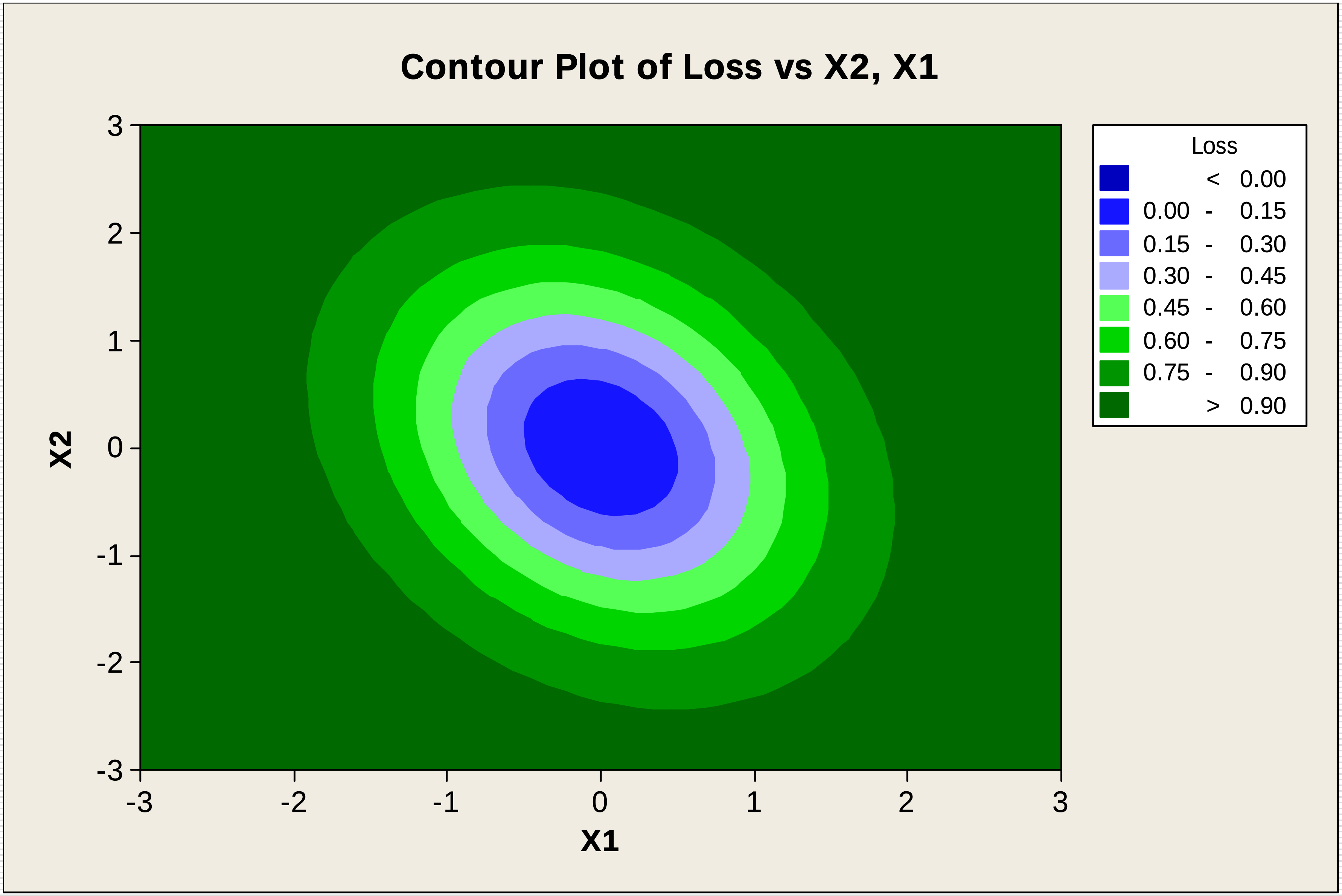
- x and τ are p-vectors
- L is a p × p positive definite matrix analogous to the covariance matrix in a multivariate normal distribution

Parameter interpretation

- Off diagonal elements express non-spherical losses:
  - Positive entries indicate antagonism: loss is greater when variables move simultaneously in the same direction
  - Negative entries indicate synergy: loss is less when variables move simultaneously in the same direction

L	0.8	-0.25
	-0.25	1.3





Contour Plot of Loss vs X2, X1
Loss
< 0.00
0.00 - 0.15
0.15 - 0.30
0.30 - 0.45
0.45 - 0.60
0.60 - 0.75
0.75 - 0.90
> 0.90
X2
3
2
1
0
-1
-2
-3
X1
-3
-2
-1
0
1
2
3

Parameter estimation

- Non-linear fitting on the basis of historical data seems the best method
- Assuming zero off-diagonal elements will probably lead to unrealistic models

Expected loss with multivariate normal process distribution

$$1-\frac{\left|L^{-1}+M^{-1}\right|^{-\frac{1}{2}}}{\left|M\right|^{-\frac{1}{2}}}e^{-\frac{1}{2}\left(\mu^T M^{-1}\mu+\tau^T L^{-1}\tau\right)}\times$$

$$e^{\frac{1}{2}\left(\mu^T M^{-1}+\tau^T L^{-1}\right)\left(L^{-1}+M^{-1}\right)\left(M^{-1}\mu+L^{-1}\tau\right)}$$

Exploiting MINLF

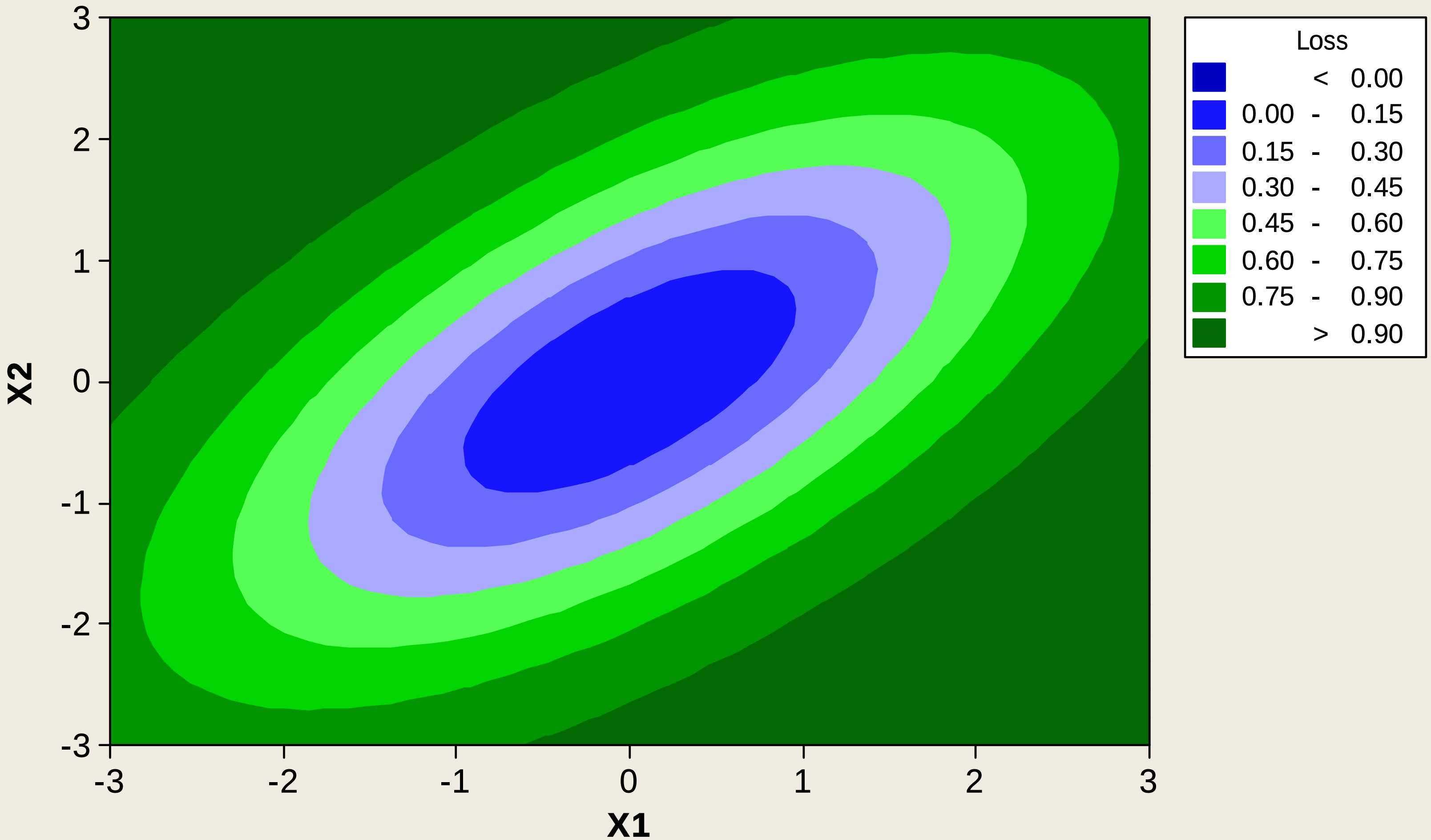
- Quantify process quality with expected loss
- Evaluate alternative process targets
- Evaluate process and equipment changes
- Predict results from feed-forward process control
- Optimize feed-forward schemes

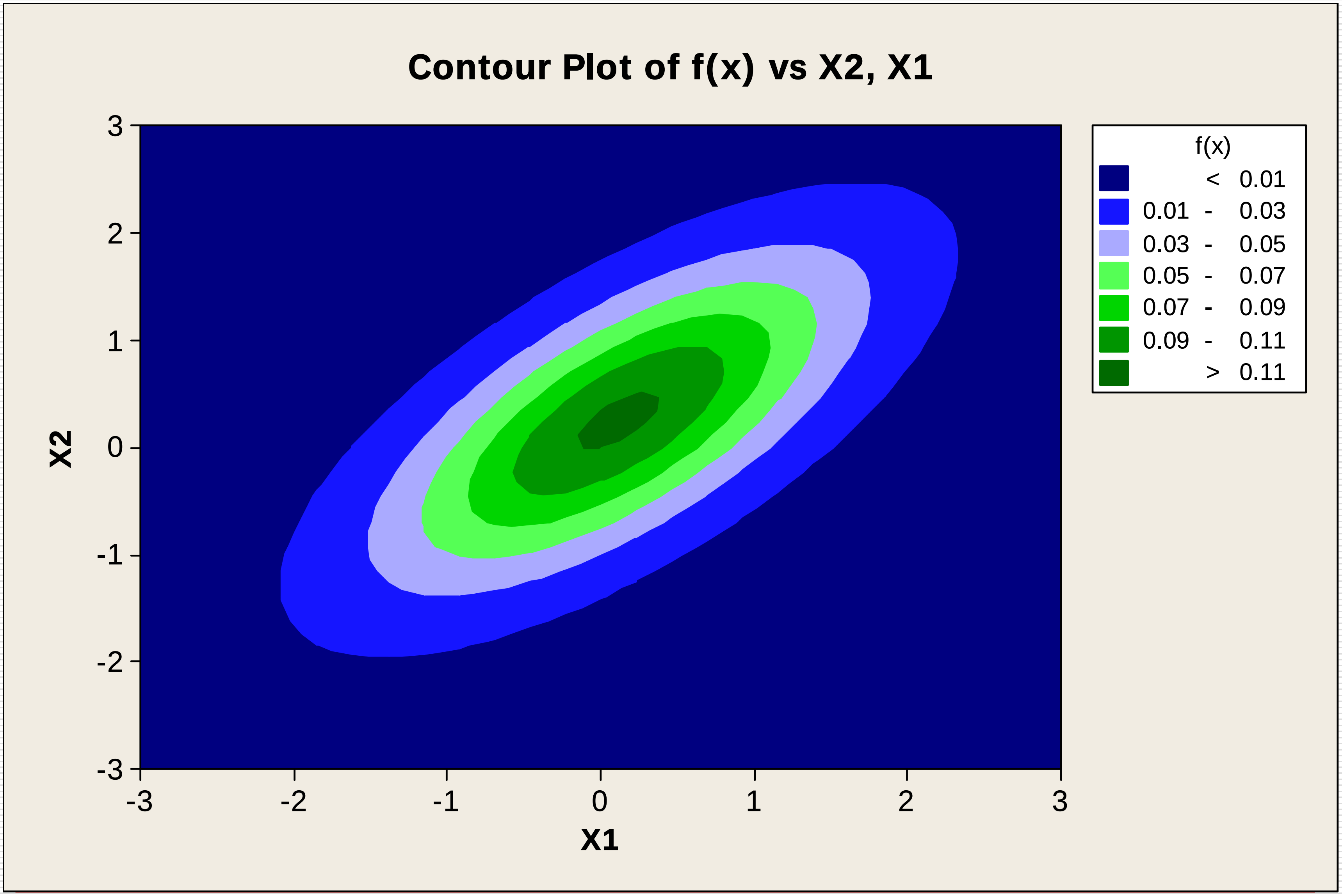
Process loss example

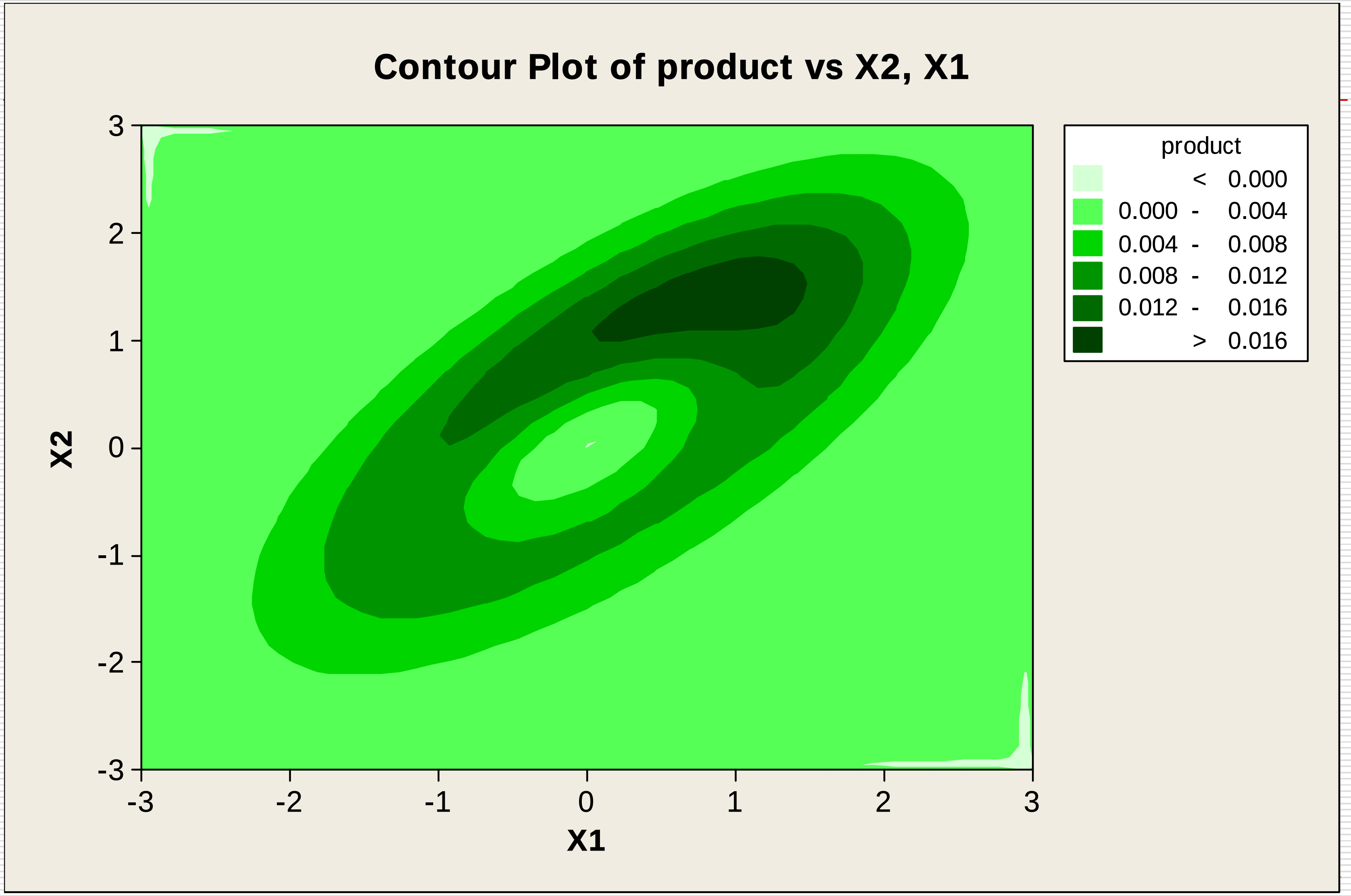
- Normally distributed process and MINLF

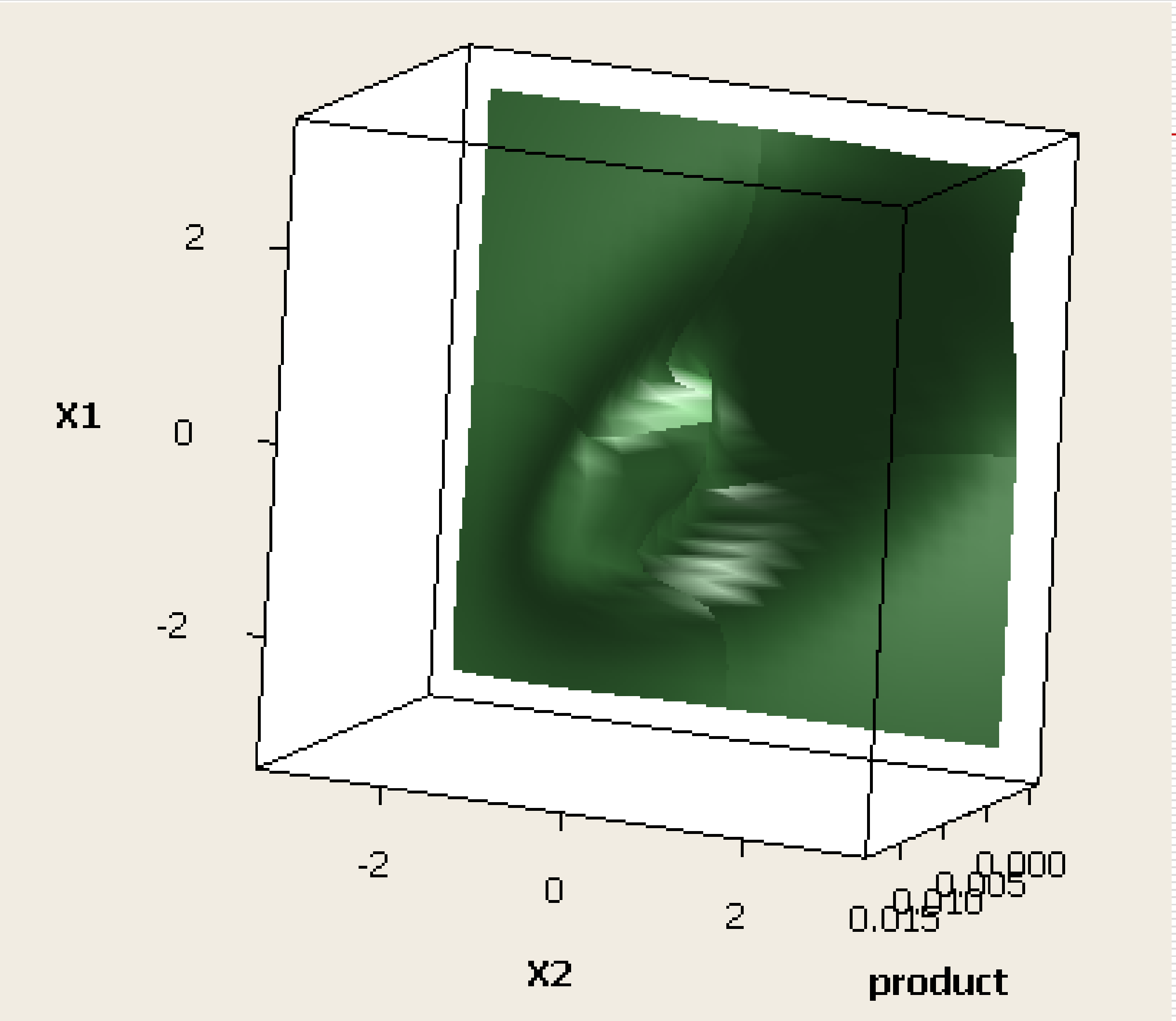
M	1.000	0.700
	0.700	1.000
L	2.89	1.802
	1.802	2.66
mu	0.12	
	0.25	
tau	0	
	0	
EL	**0.261**	

Contour Plot of Loss vs X2, X1

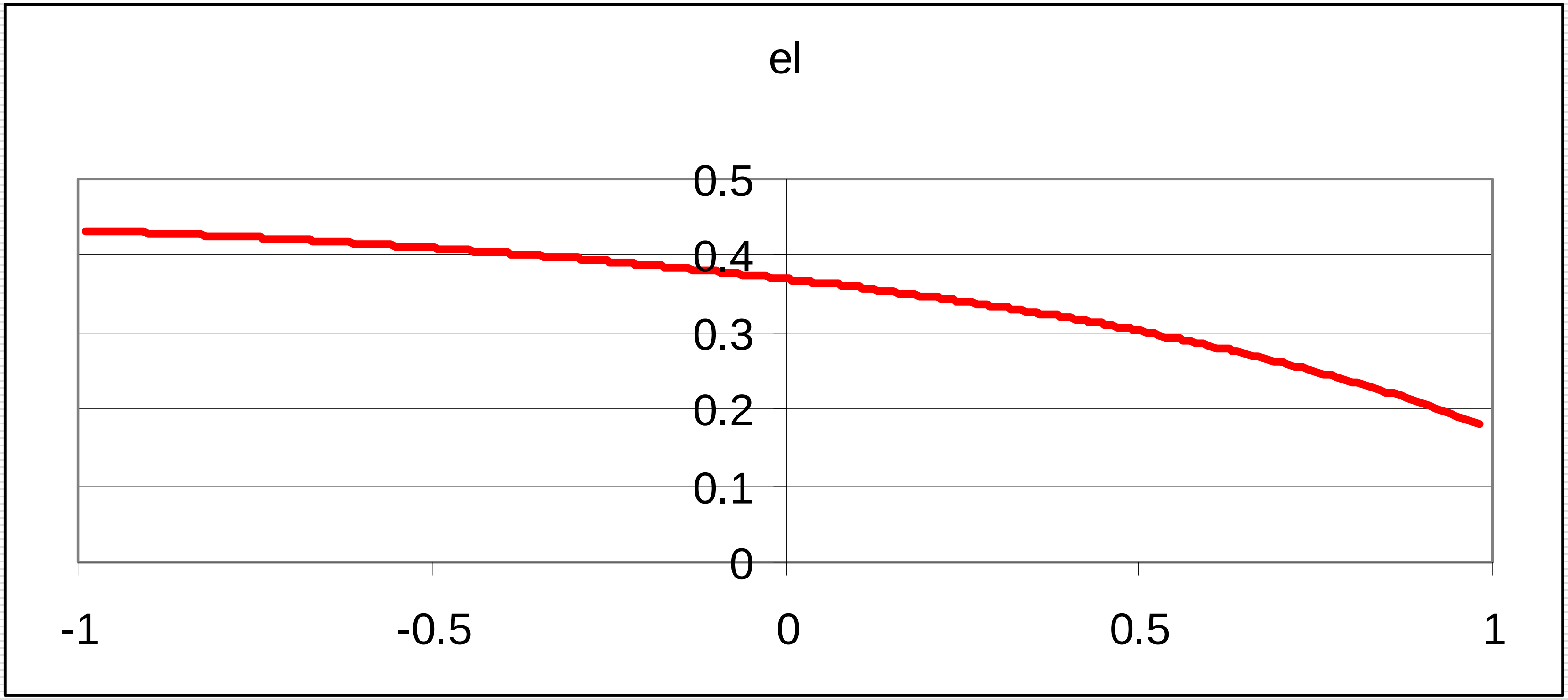



Contour Plot of f(x) vs X2, X1
X2
X1
3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3
f(x)
< 0.01
0.01 - 0.03
0.03 - 0.05
0.05 - 0.07
0.07 - 0.09
0.09 - 0.11
> 0.11


Contour Plot of product vs X2, X1
product
< 0.000
0.000 - 0.004
0.004 - 0.008
0.008 - 0.012
0.012 - 0.016
> 0.016
X2
3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3
X1


2
X1
0
-2
-2
0
2
X2
0.000
0.005
0.010
0.015
product

Effect of process correlation



Ongoing and future research

- Creating a “replacement package” for Taguchi loss functions
  - Naresh Sharma, Beth Cudney
- Documenting our research in robust regression with INLF
  - Lance Kaminski
- Push the case for replacement of Cpk as a process health indicator
  - Melissa Baeten

References

- D.C. Drain and A.M. Gough, “Applications of the Upside-Down Normal Loss Functions, “IEEE Transactions on Semiconductor Manfuacturing, Vol. 9 No. 1, pp 143-145, 1996.
- B.P.K. Leung and Fred Spiring, “Some Properties of the Family of Inverted Probability Loss Functions”, *Quality Technology & Quantitative Management*, Vol1, No 1, pp 125-147, 2004.