# Medelvärdesbildning – exakt hur mycket hjälper det?

*av Mattias Ericsson*

En av de allra vanligaste sätten att förbättra sin mätprocess är att ta medelvärdet på ett antal mätningar. Av erfarenhet vet vi att det nästan alltid blir bättre mätningar med mindre variation, men vet samtidigt att vi inte får medelbilda över ett för långt tidsintervall att vi riskerar att missa förändringar i signalen som vi faktiskt är intresserade av. Dessa två motstridiga faktorer måste tas hänsyn till.

Frågan vi ställer oss är, hur mycket minskar variationen i signalen med hjälp av medelvärdesbildning? Om vi börjar med att definiera variation, så är ett mycket vanligt mått *standardavvikelse,* σ. Detta är ett mått på avvikelsen från *medelvärdet*, µ. Ju fler mätpunkter som ligger långt ifrån medelvärdet, desto större standard avvikelse. Standardavvikelsen har samma måttenhet som själva mätvärdet, vilket gör det enkelt att använda och förstå.

$$σ=\sqrt{\frac{\sum\_{i=1}^{n}\left(x\_{i}-\overbar{x}\right)^{2}}{n-1}}$$

De allra flesta mätprocesser är *normalfördelade*, vilket innebär att de är symetriskt centrerade kring medelvärdet i den berömda klockformen. När en mätprocess är normalfördelad, kan man anta att ca 68% av alla mätvärden ligger inom ± 1 standardavvikelse från medelvärdet och >99% inom ± 3 standardavvikelser från medelvärdet. Se figur nedan på normalfördelningskurvan. Arean under kurvan är 100%.



Av erfarenhet vet vi att variationen i en mätprocess minskar när vi medelvärdesbildar mätningen och detta måste då följaktligen innebära att standardavvikelsen också minskar på de medelvärdesbildade grupperna av mätvärden, men hur mycket? Enklast är att visa med ett enkelt exempel:

Anta att vi har en mätning av en 5V spänning som ser ut enligt figuren nedan.



Här har vi 50000 mätpunkter av en 5V spänning. Vi kan enkelt beräkna medelvärdet till 5.00V och standardavvikelsen (enligt formeln ovan) till 0.099642V. Detta innebär att >99% av alla mätvärden ligger 5.00±0.298926V. Detta är en lite för stor variation för att vi ska vara nöjda, så vi beslutar oss för att medelvärdesbilda mätningen i grupper med 100 mätpunkter i taget och vi får då nya mätningar med medelvärdet av grupperna enligt nedan.



Vi kan direkt se att variationen nu är betydligt mindre och vi beräknar återigen medelvärdet och standardavvikelsen för våra gruppers medelvärden. Medelvärdet av grupperna beräknas till 5.00V och standardavvikelsen till 0.009996V. Detta innebär att >99% av alla mätvärden nu ligger inom 5.00±0.029988V. Detta är fortfarande inte riktigt så bra som vi vill nå, så vi testar att medelvärdesbilda över grupper av 500 mätpunkter istället och får nya mätningar enligt nedan.



Vi beräknar återigen medelvärdet och standardavvikelsen på de nya grupperna till 5.00V respektive 0.004342V. Detta innebär att >99% av alla mätvärden kommer att ligga inom 5.00±0.01303V, vilket vi denna gång anser vara godkänt.

Nu kan några intressanta iakttagelser göras:

1. Medelvärdet är oförändrat, vilket är ganska logiskt. Medelvärdet av alla mätpunkter är samma som medelvärdet av grupperna av medelvärden.
2. Standardavvikelsen blir mindre och mindre ju fler mätpunkter vi använt i varje grupp vid medelvärdesbildning och om vi beräknar kvoten mellan den ursprungliga standardavvikelsen och de två medelvärdesbildade mätningar får vi följande kvoter:
	1. $\frac{0.099642}{0.009996}≈9.97$
	2. $\frac{0.099642}{0.004342}≈22.9$

Trots att vi ökade medelvärdesbildningen fem gånger från 100 mätpunkter till 500 mätpunkter, blev inte skillnaden fem gånger större. Sambandet mellan variationen av de medelvärdesbildande grupperna och de ursprungliga mätvärdena definieras av *centrala gränsvärdessatsen,* vilken anger sambandet till $σ\_{\overbar{x}}=\frac{σ}{\sqrt{n}}$, där n är antal mätpunkter i medelvärdesbildningen. Låt oss testa teorin:

1. $\sqrt{100}=10.0$
2. $\sqrt{500}=22.4$

Detta är ganska exakt samma sak som vi faktiskt kan konstatera rent praktiskt. Ju fler grupper av medelvärden vi använder oss av och beräknar standardavvikelsen över, ju mer kommer den att närma sig det teoretiska.

Vi får det enkla sambandet att variationen minskar med $\sqrt{n}$, där antal mätpunkter som vi använder vid medelvärdesbildningen ger oss följande tumregler.

Medelvärdesbildning ger följande förbättringar:

|  |  |
| --- | --- |
| Antal mätpunkter, n | Minskad variation, $\sqrt{n}$ (ggr) |
| 2 | 1.4 |
| 10 | 3.2 |
| 100 | 10 |
| 1000 | 32 |
| 10000 | 100 |
| 100000 | 316 |

Slutsatsen är att det hjälper inte speciellt mycket att öka medelvärdesbildningen, används >1000 mätpunkter blir skillnaden inte så stor om vi ökar till 10000, däremot förloras mycket information i tidsplanet och det tar dessutom lång tid att mäta.

Referenser:

Bass, Issa, 2007: Six Sigma Statistics with Excel and Minitab. ISBN: 978-0-07-149646-9.

Om författaren:

*Mattias Ericsson har jobbat med test- och mätsystem i 15 år. Han är R&D manager inom AddQ Göteborg och jobbar med framtagningen av AddQs resultathanteringssystmet QRM när han inte är upptagen med kundprojekt.*