Geometric and Arithmetic means

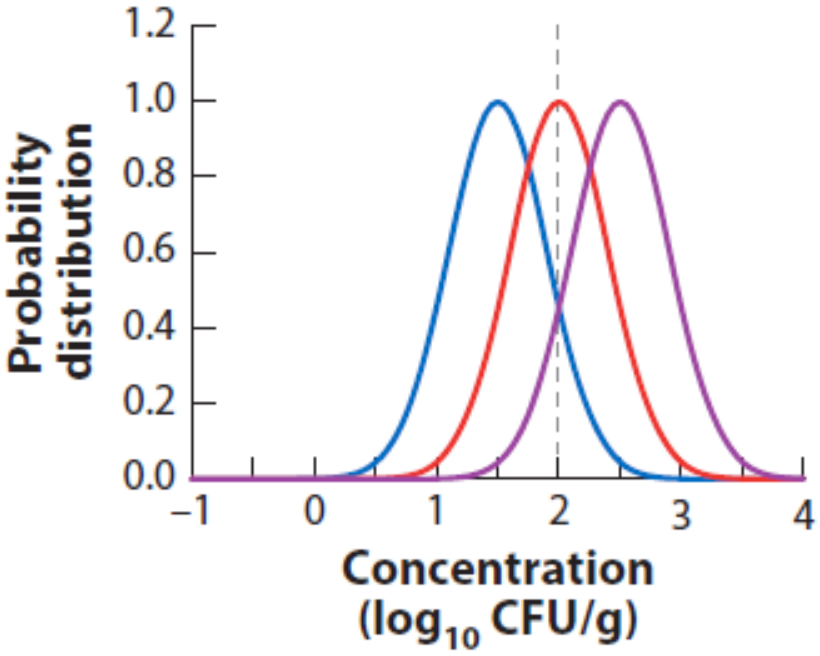
Marcel Zwietering

Wageningen University

ICMSF Member since 2005

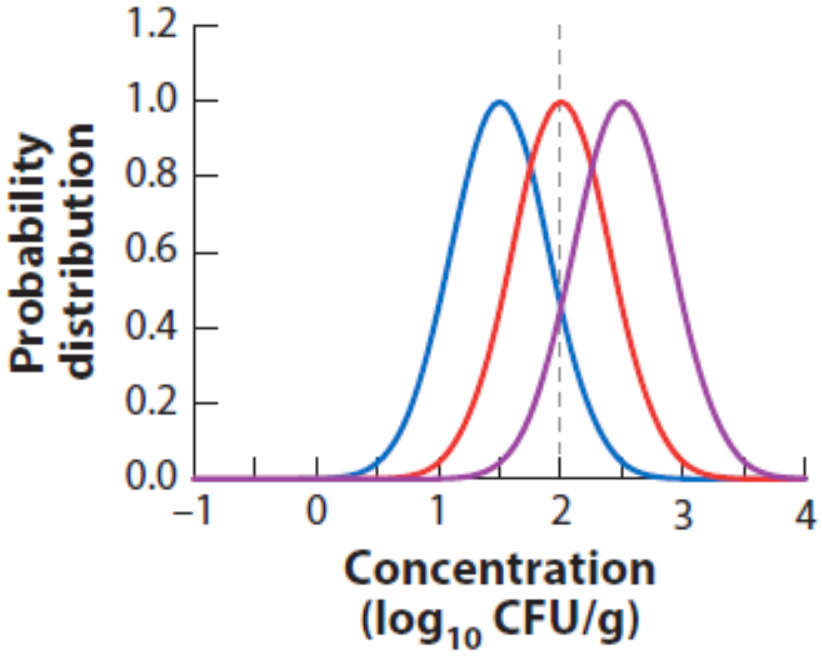
WAGENINGEN
UNIVERSITY & RESEARCH

Contamination often log-normally distributed



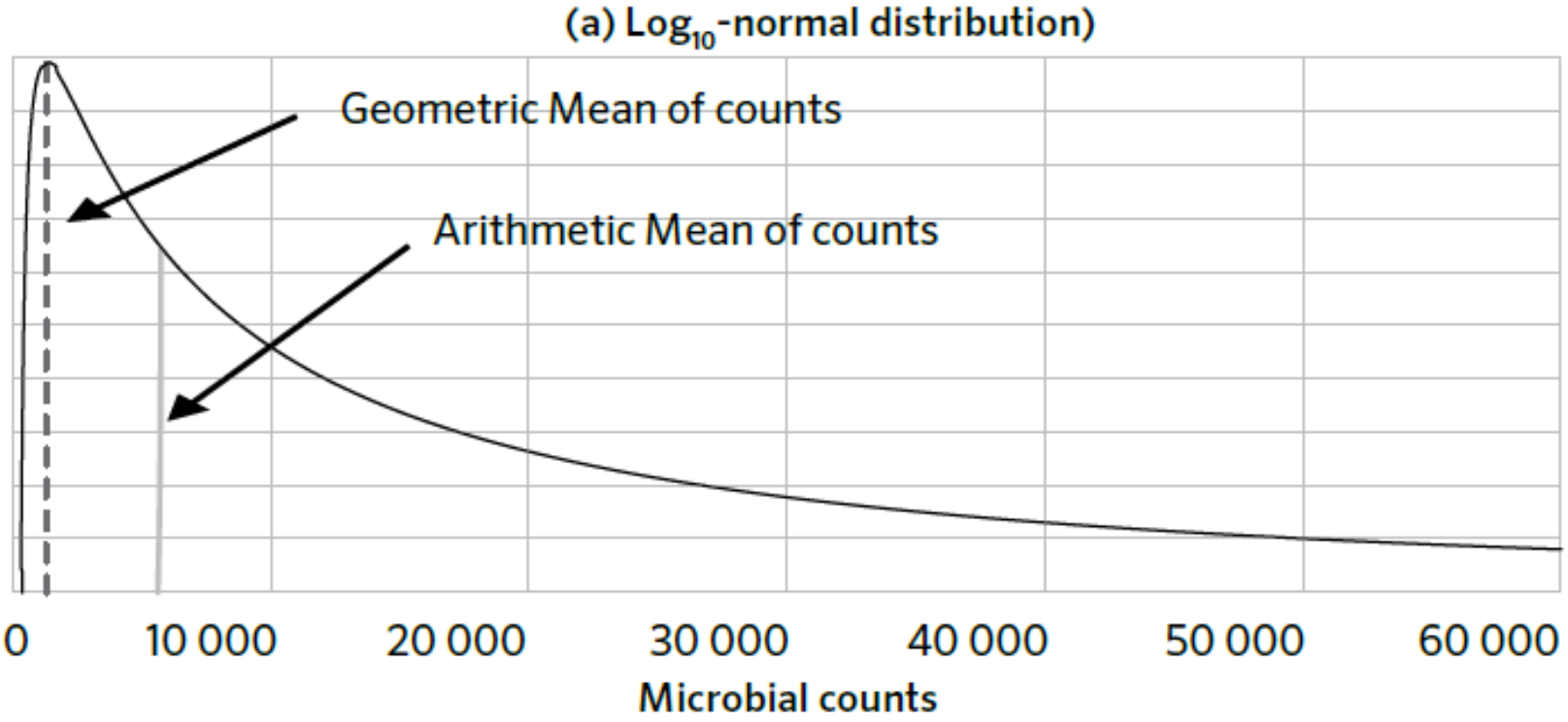
Define the mean on log-scale or on arithmetic scale ?

Contamination often log-normally distributed



Geometric mean: Concentration at the peak of the distribution

Arithmetic mean: Average on the linear scale



Geometric mean: Concentration at the peak of the distribution

Arithmetic mean: real average on the linear scale (more influenced by the right hand tail with big numbers)

Which one to chose ?

For microbial analysis for data representation: log scale

For impact analysis: linear scale

You have to eat 3 products that contain *B. cereus*

Your choices:

1: 100 cfu	100 cfu	100 cfu
2: 10 cfu	100 cfu	1000 cfu
3: 1 cfu	100 cfu	10000 cfu

Which row would you chose ?

You have to eat 3 products that contain *B. cereus*

Your choices:

1: 100 cfu	100 cfu	100 cfu
2: 10 cfu	100 cfu	1000 cfu
3: 1 cfu	100 cfu	10000 cfu

Log of the concentration

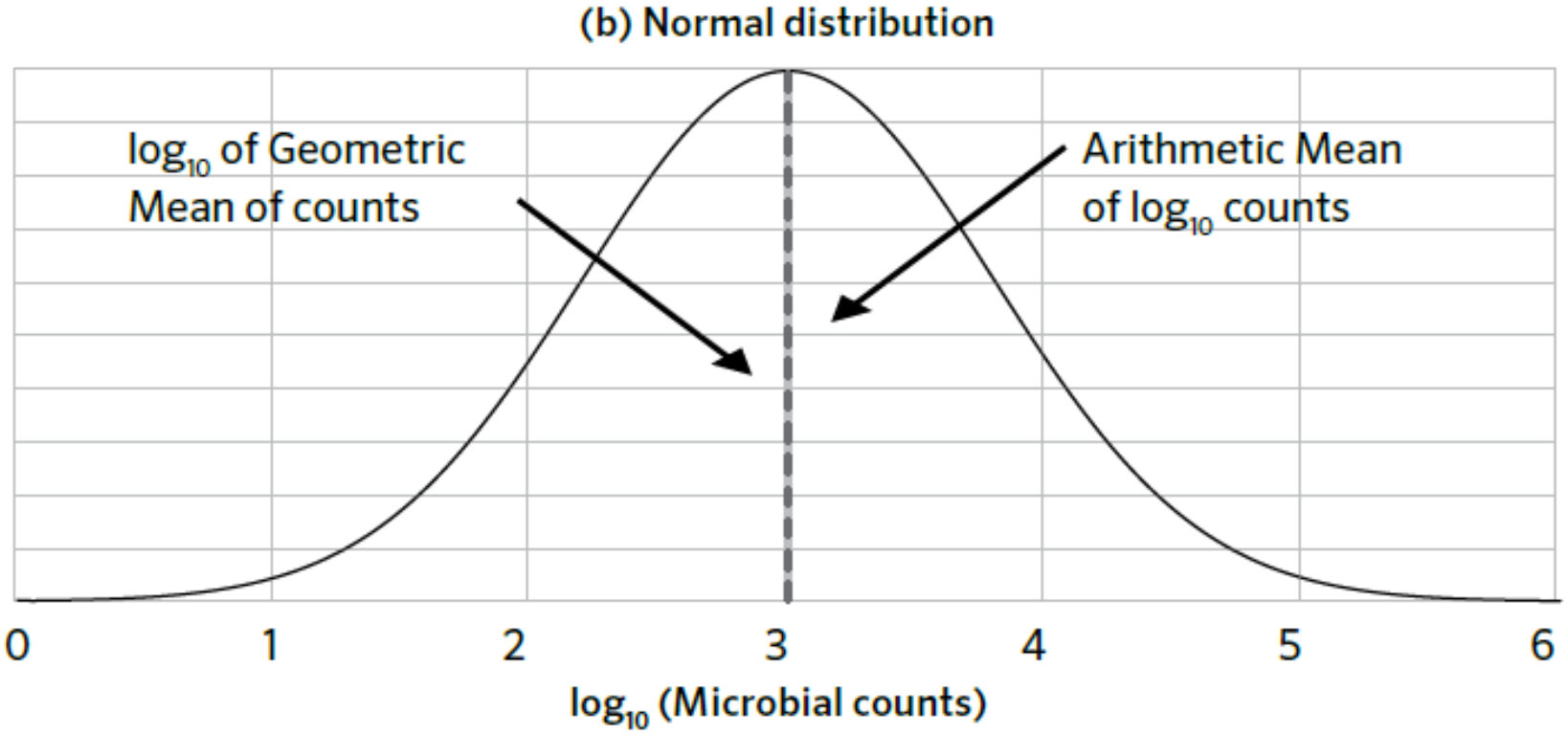
1:	2	2	2: mean is 2
2:	1	2	3: mean is 2
3:	0	2	4: mean is 2

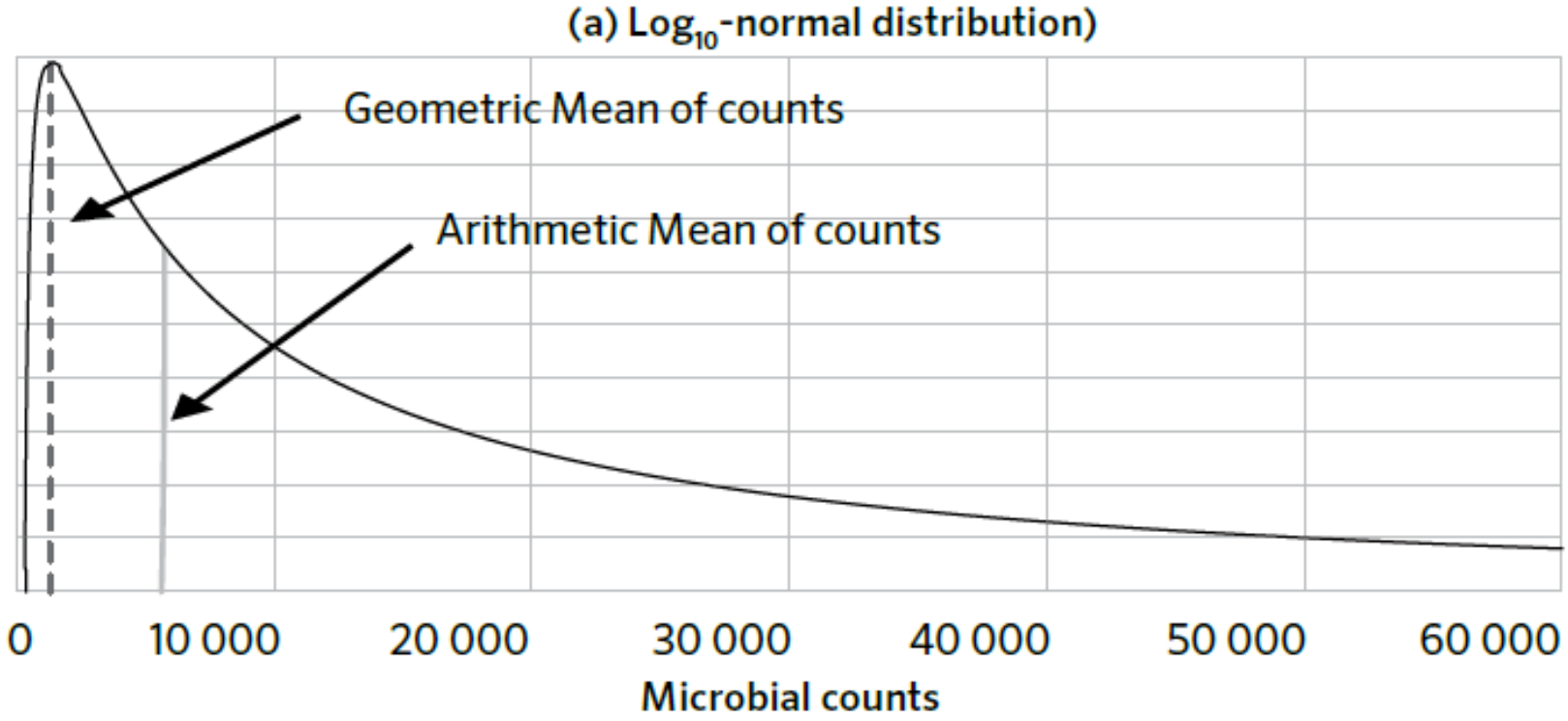
You have to eat 3 products that contain *B. cereus*

Your choices:

1: 100 cfu	100 cfu	100 cfu
2: 10 cfu	100 cfu	1000 cfu
3: 1 cfu	100 cfu	10000 cfu

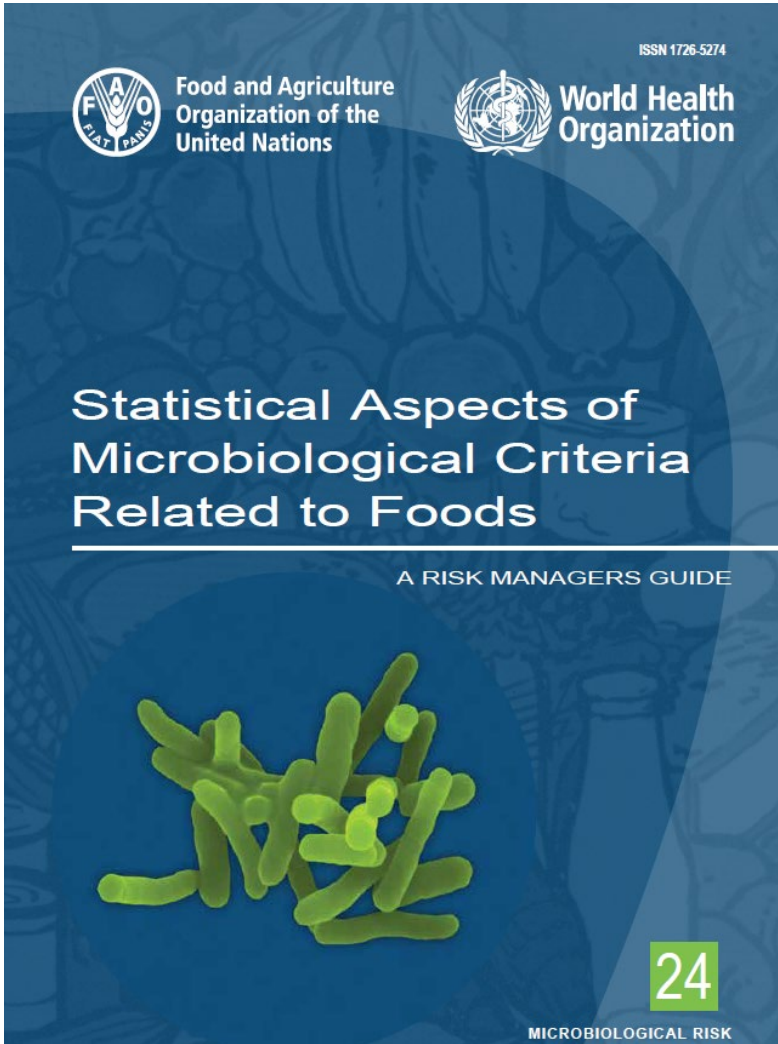
	arithmetic mean	log(arithmetic mean)
1:	100 cfu/g	2.0
2:	370 cfu/g	2.6
3:	3367 cfu/g	3.5


(b) Normal distribution
log10 of Geometric Mean of counts
Arithmetic Mean of log10 counts
0
1
2
3
4
5
6
log10 (Microbial counts)


(a) Log10-normal distribution)
Geometric Mean of counts
Arithmetic Mean of counts
0
10 000
20 000
30 000
40 000
50 000
60 000
Microbial counts

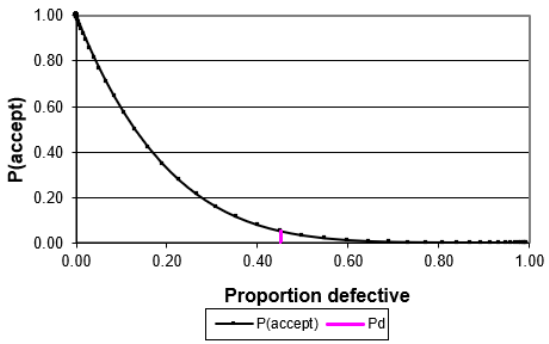
For a lognormal distribution

$$\log_{10}(\bar{C}) = \overline{\log_{10}(C)} + 0.5 \cdot \sigma^2_{\log_{10}(C)}$$

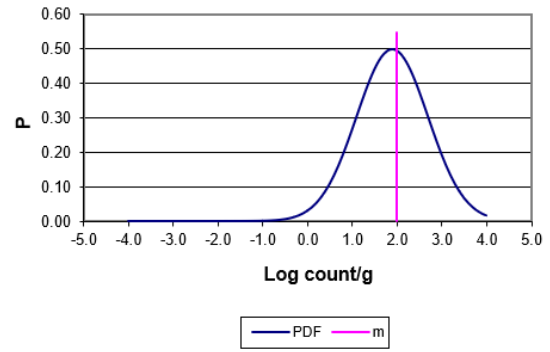


m | 2

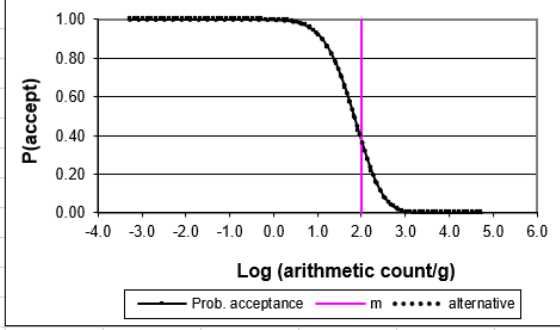
Operating characteristic curve for proportion defective, with n=5 and c=0



Probability density function (PDF) for log counts. Distribution mean = 1.90 and sigma = 0.80



Operating characteristic curve scaled to relate log arithmetic mean count to m



Lot acceptance for Pd			P(accept)	
Pd	10	%	59.0	%
actualPd	45.1	%	5.00	%

INPUTS		P(accept)		
mean	1.90	Computed	5.00	%
sigma	0.80	Desired	5	%
m	2			
n	5			
c	0			

Find mean that gives desired P(accept)

Find n that gives desired P(accept) or better (less)

P(reject) 95.00 %

ALTERNATIVE n AND c		P(accept)		
mean	1.90	Computed	5.00	%
sigma	0.80	Target, left	5.00	%
m	2			
n	5			
c	0			

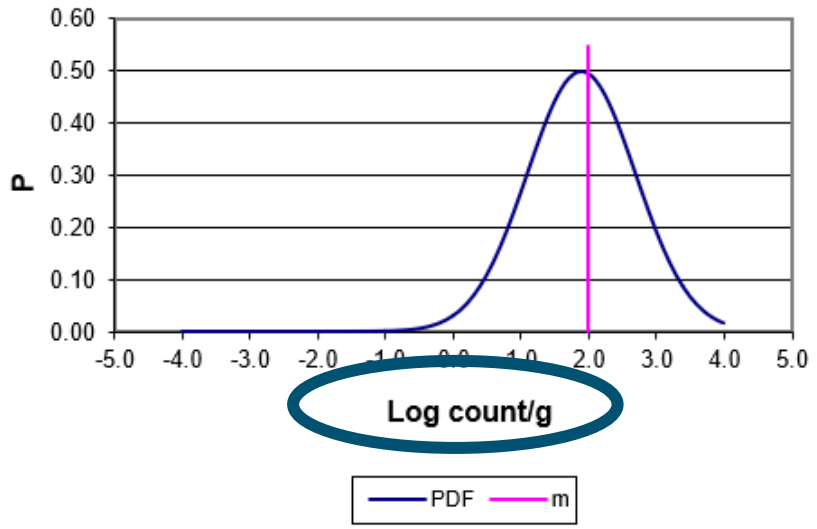
For any value of n and c imputed find the m that gives the same P(accept) as the model on the

Sandbox: for your own calculations			
FSO	3		
complient	0.91526	not comply	0.08474
%	91.52604	%	8.473957
	1.339231		
	1.984384		

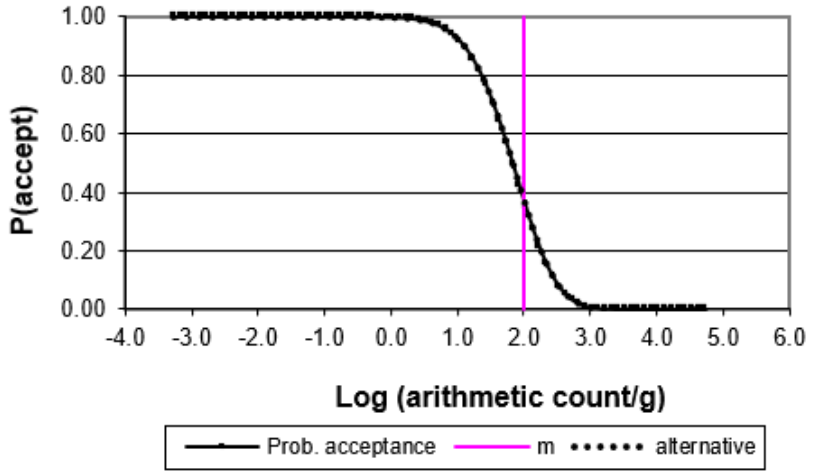
Means			
Arithmetic (=Average)		Geometric (=Median)	
434.2	cfu/g	79.6	cfu/g
2.64	log cfu/g	1.90	log cfu/g

Implied Acceptance level		
Percentile	z-score	Concentration at this percentile
99.9	3.10	4.38
FSO	P(exceedance)	target level
4.38	0.000973	1.900897

Probability density function (PDF) for log counts. Distribution mean = 1.90 and sigma = 0.80



Operating characteristic curve scaled to relate log arithmetic mean count to m



INPUTS	
mean	1.90
sigma	0.80
m	2
n	5
c	0

P(accept)		
Computed	5.00	%
Desired	5	%

Find mean that gives desired P(accept)

Find n that gives desired P(accept) or better (less)

P(reject)	95.00	%

ALTERNATIVE n AND c	
mean	1.90
sigma	0.80
m	2
n	5
c	0

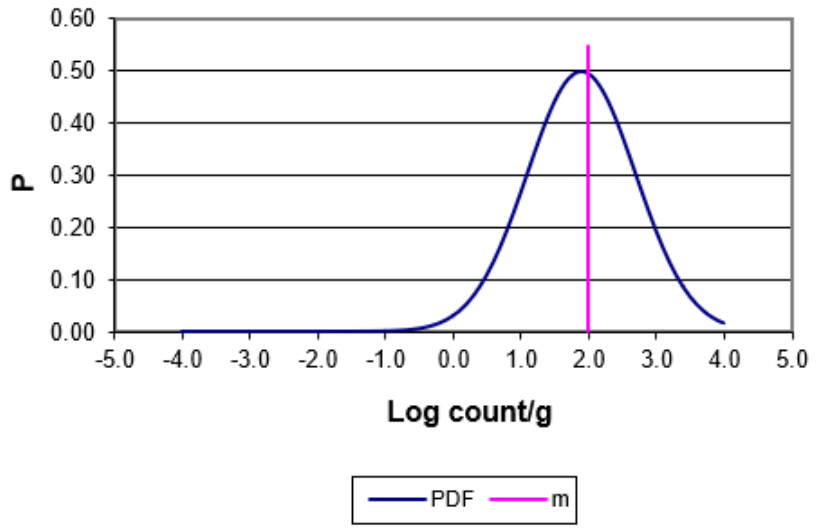
P(accept)		
Computed	5.00	%
Target, left	5.00	%

For any value of n and c imputed find the m that gives the same P(accept) as the model on the

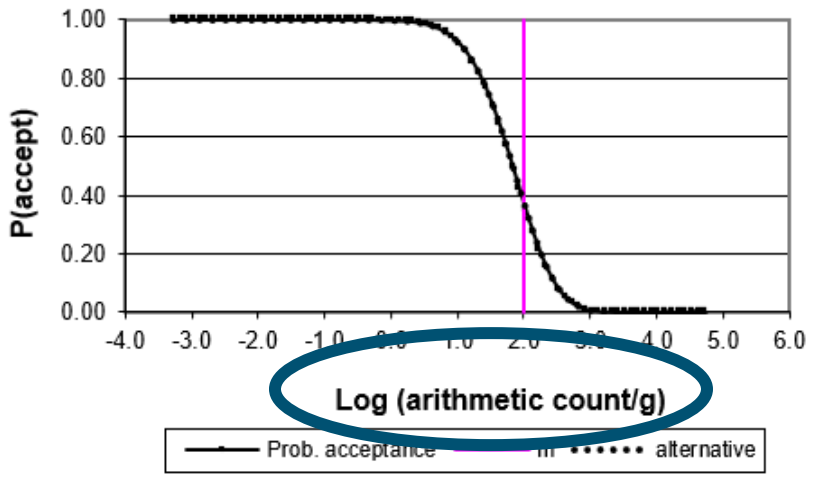
Means			
Arithmetic (=Average)		Geometric (=Median)	
434.2	cfu/g	79.6	cfu/g
2.64	log cfu/g	1.90	log cfu/g

Implied Acceptance level		
Percentile	z-score	Concentration at this percentile
99.9	3.10	4.38
FSO	P(exceedance)	target level
4.38	0.000973	1.900897

Probability density function (PDF) for log counts. Distribution mean = 1.90 and sigma = 0.80



Operating characteristic curve scaled to relate log arithmetic mean count to m



INPUTS	
mean	1.90
sigma	0.80
m	2
n	5
c	0

P(accept)		
Computed	5.00	%
Desired	5	%

Find mean that gives desired P(accept)

Find n that gives desired P(accept) or better (less)

P(reject)	95.00	%

ALTERNATIVE n AND c	
mean	1.90
sigma	0.80
m	2
n	5
c	0

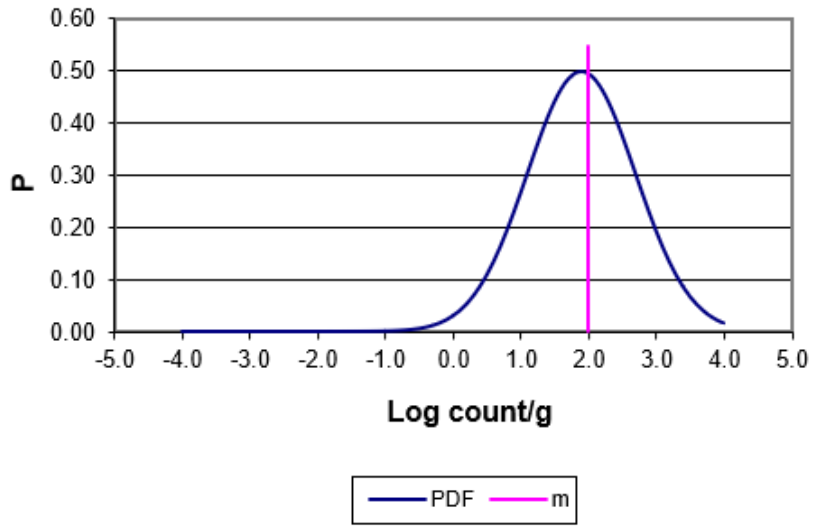
P(accept)		
Computed	5.00	%
Target, left	5.00	%

For any value of n and c imputed find the m that gives the same P(accept) as the model on the

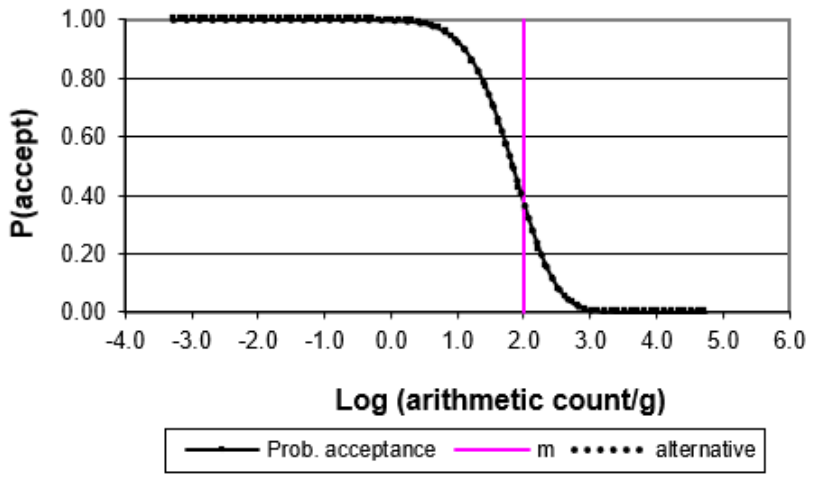
Means			
Arithmetic (=Average)		Geometric (=Median)	
434.2	cfu/g	79.6	cfu/g
2.64	log cfu/g	1.90	log cfu/g

Implied Acceptance level		
Percentile	z-score	Concentration at this percentile
99.9	3.10	4.38
FSO	P(exceedance)	target level
4.38	0.000973	1.900897

Probability density function (PDF) for log counts. Distribution mean = 1.90 and sigma = 0.80



Operating characteristic curve scaled to relate log arithmetic mean count to m



INPUTS	
mean	1.90
sigma	0.80
m	2
n	5
c	0

P(accept)		
Computed	5.00	%
Desired	5	%

Find mean that gives desired P(accept)

Find n that gives desired P(accept) or better (less)

P(reject) 95.00 %

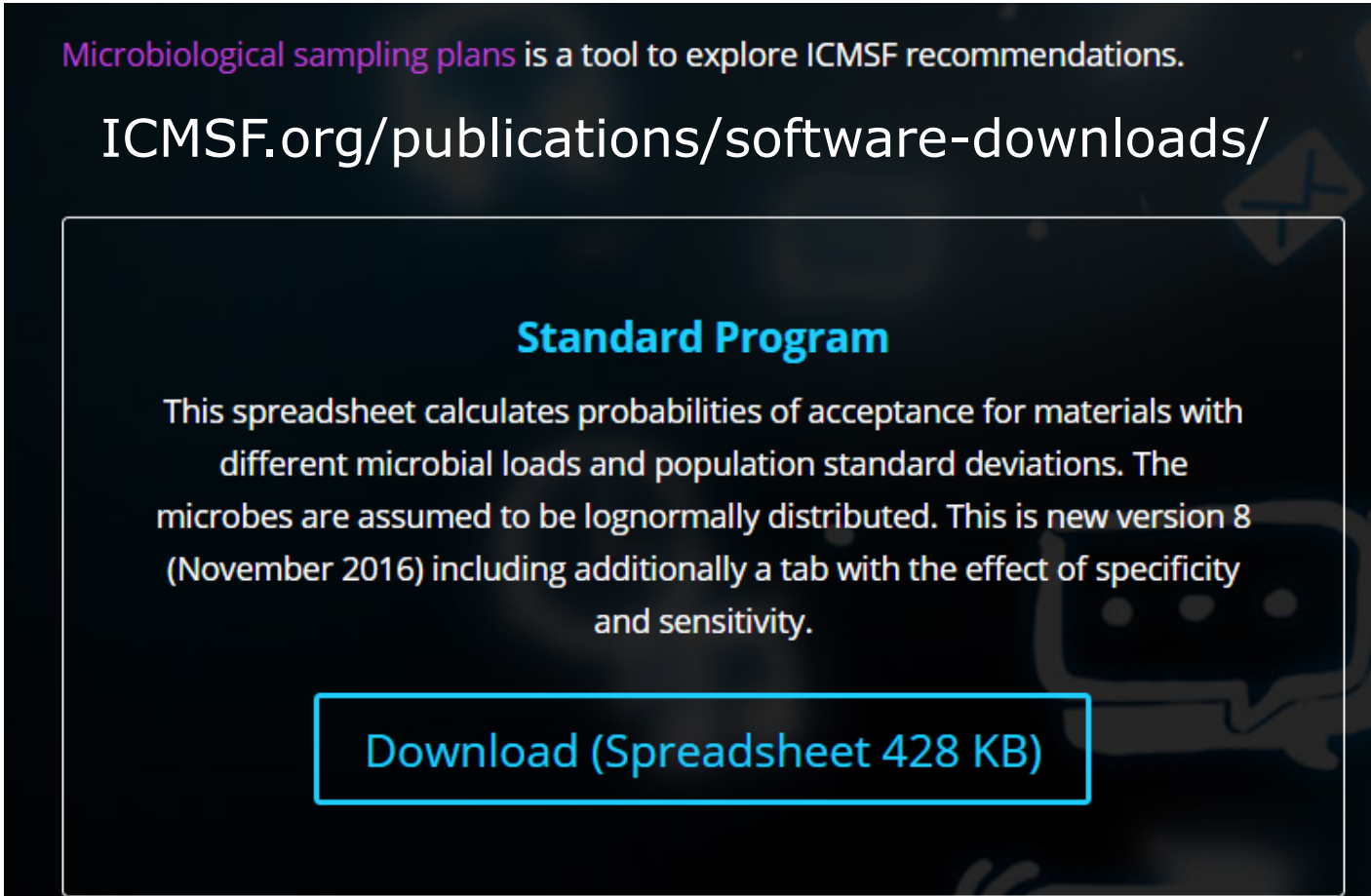
ALTERNATIVE n AND c	
mean	1.90
sigma	0.80
m	2
n	5
c	0

P(accept)		
Computed	5.00	%
Target, left	5.00	%

For any value of n and c imputed find the m that gives the same P(accept) as the model on the

Means			
Arithmetic (=Average)		Geometric (=Median)	
434.2	cfu/g	79.6	cfu/g
2.64	log cfu/g	1.90	log cfu/g

Implied Acceptance level		
Percentile	z-score	Concentration at this percentile
99.9	3.10	4.38
FSO	P(exceedance)	target level
4.38	0.000973	1.900897

Conclusions

- Distributions relevant for performance of sampling plans
- Arithmetic mean is most relevant for risk management
- Tools exist !