

---

# Approximating the Sum of Lognormal Distributions to Enhance Models of Inhalational Anthrax

---

William R. Hogan, MD, MS

Garrick L. Wallstrom, PhD

Department of Biomedical Informatics

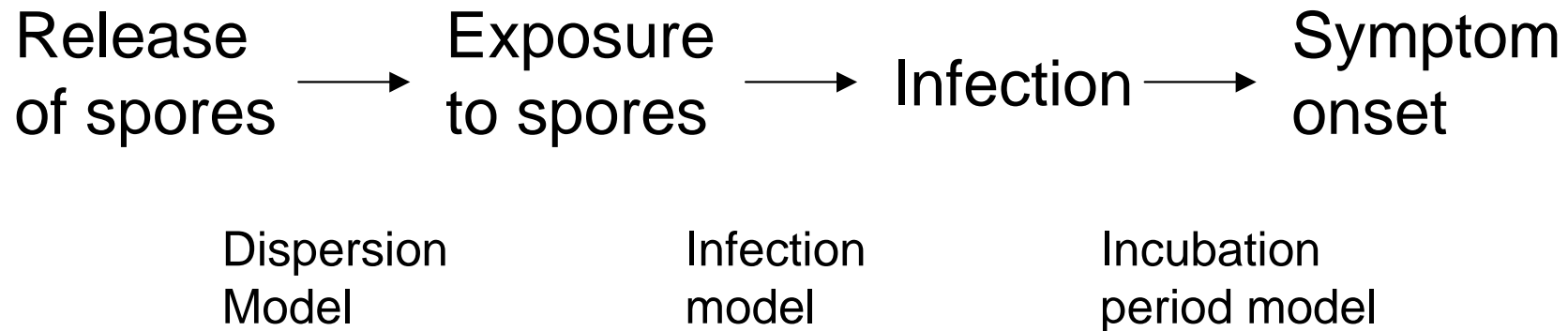
University of Pittsburgh

---

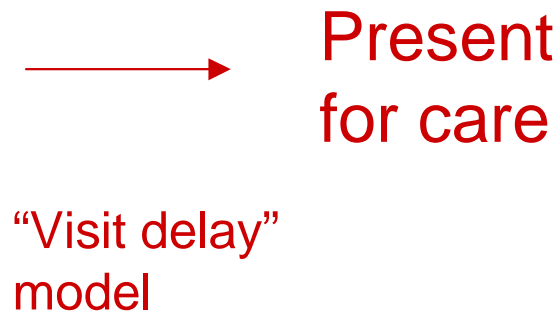
# Outline

- Motivation
  - Approach
  - Application
  - Conclusion
-

# Motivation



GOAL IS TO ADD:



# Example

- What is the probability  $p$  that an individual who inhaled  $d$  spores will visit an emergency department in a given 24-hour interval?

$$P(\text{visit ED at } t_i < t < t_{i+24} \mid d, I) = \int_{t_i}^{t_{i+24}} \int_0^{t_{i+24}} f_S(t_S \mid d, I) f_V(t_V \mid I) dt_S dt_V$$

- Where  $f_S(\cdot)$ ,  $f_V(\cdot)$  are lognormal pdf of incubation period and visit delay, respectively, and  $I$ =infected
-

---

# Approach

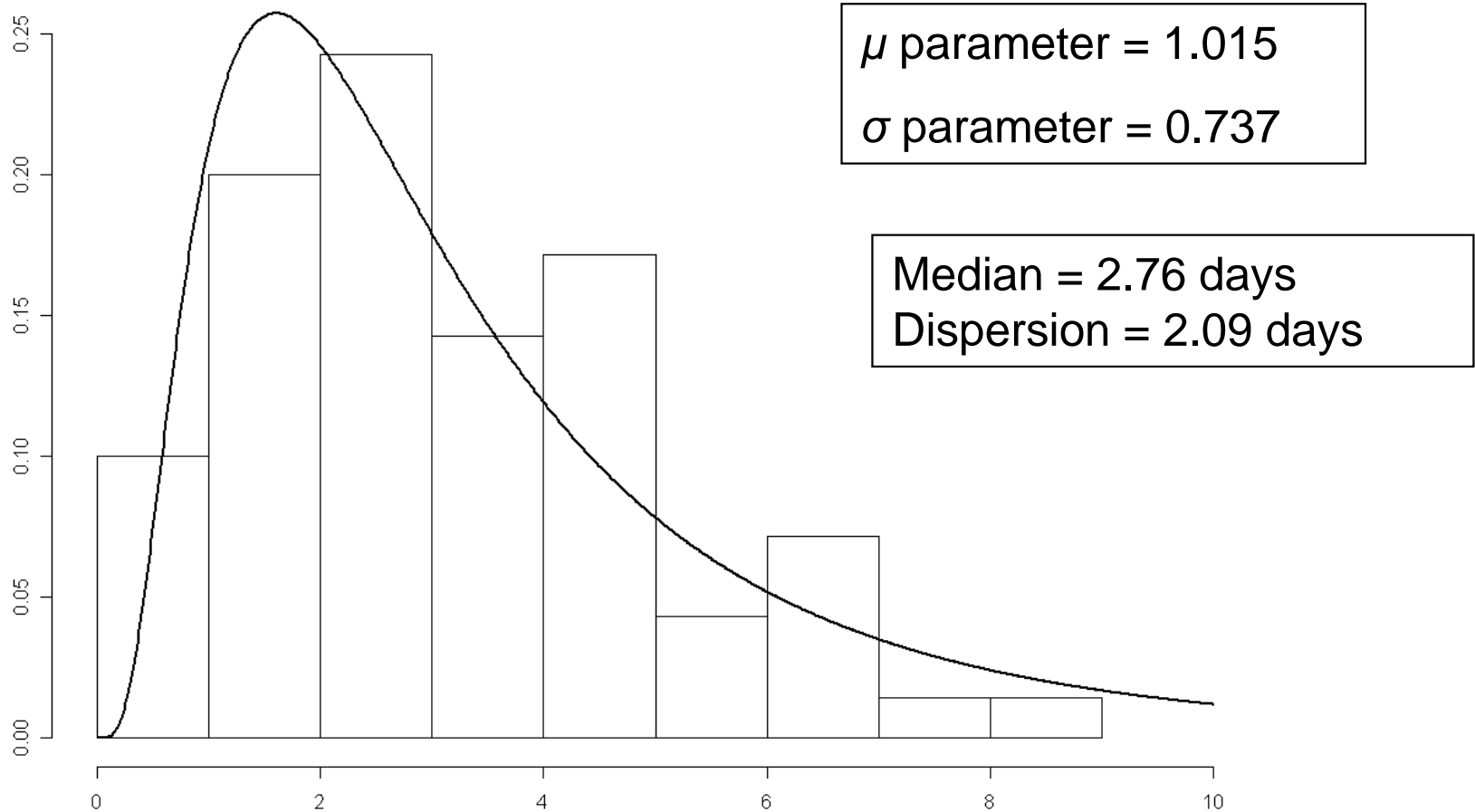
- Model visit delay as lognormal distribution, using maximum-likelihood estimation to derive parameters from data
- Approximate the sum of the two lognormal distributions with a single lognormal, then compute the AUC between  $t_i$  and  $t_{i+24}$  under its pdf:

$$P(ED \text{ at } t_i < \underbrace{t_{incub} + t_{visit}} < t_{i+24} \mid d, I)$$

Approximate this sum of two lognormals  
as a single lognormal distribution!

---

# Visit Delay: Estimation from Data



# Approximation of Sum of Lognormals

- Method is from Wu et al, 2005
- Assume distributions are independent
- Thus the moment-generating function (MGF) of their sum is the product of their individual MGFs (MGF of  $X$  is  $E(e^{tX})$ )
- Lognormal MGF:

$$\Psi(t) = \int_0^{\infty} \frac{\exp(tx)}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right] dx$$

# Approximation Method (continued)

- Using Gauss-Hermite approximation, lognormal MGF is:

$$\hat{\Psi}(t) = \sum_{n=1}^N \frac{w_n}{\sqrt{\pi}} \exp\left[t \exp\left(\sqrt{2} \sigma a_n + \mu\right)\right]$$

- MGF of sum of lognormal variables  $X$  and  $Y$  is then approximated by:

$$\sum_{n=1}^N \frac{w_n}{\sqrt{\pi}} \exp\left[t \exp\left(\sqrt{2} \sigma a_n + \mu\right)\right] = \hat{\Psi}_X(t) \hat{\Psi}_Y(t)$$

- For any  $t$ , we can calculate the right-hand side, and there are two unknowns on the left-hand side ( $\mu$ ,  $\sigma$ )
  - Thus by using two values of  $t$ , we set up a system of two equations in two unknowns, which we solve numerically
-



# Application to Incubation Period and Visit Delay

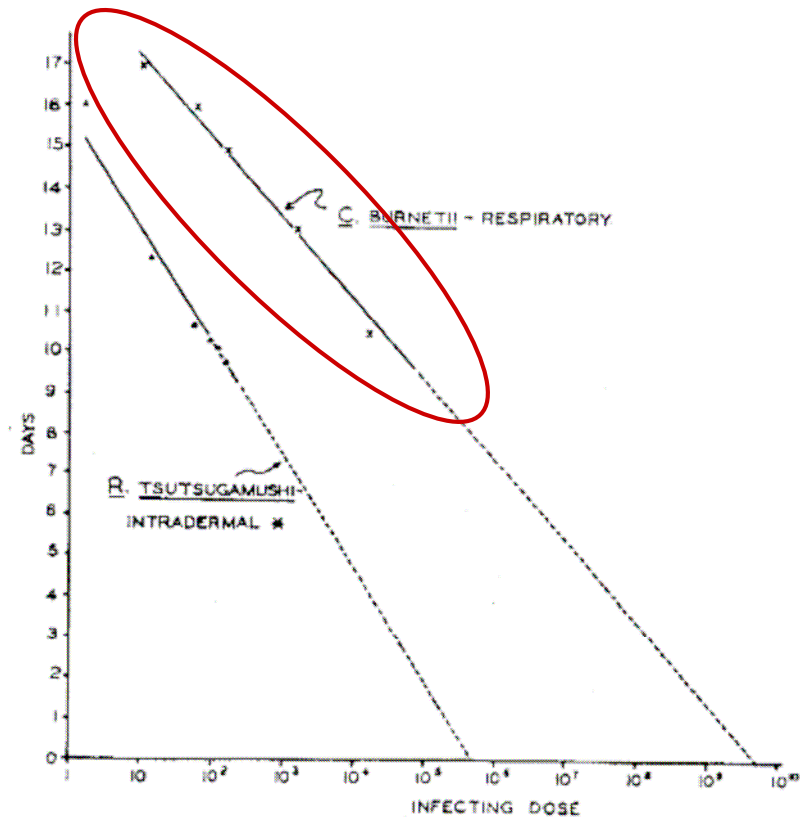
- We used  $N=5$ ,  $t_1 = -0.1$ ,  $t_2=-0.5$
- We used Wilkening's A1 model of inhalational anthrax\*:
  - ID50 = 8600 spores
  - Probit slope = 0.67
  - Relationship b/w dose/incubation period:
    - $\mu = \log(10.3 - 1.35*\log_{10}(\text{dose}))$
    - $\sigma = 0.804 - 0.079*\log_{10}(\text{dose})$
- We applied the method to lognormals for visit delay and for incubation period for  $\log_{10}(\text{dose}) = 1, 2, 3, 4, 5, 6, 7$

---

\* Wilkening DA. Sverdlovsk revisited: modeling human inhalation anthrax. Proc Natl Acad Sci U S A 2006;103(20):7589-94.

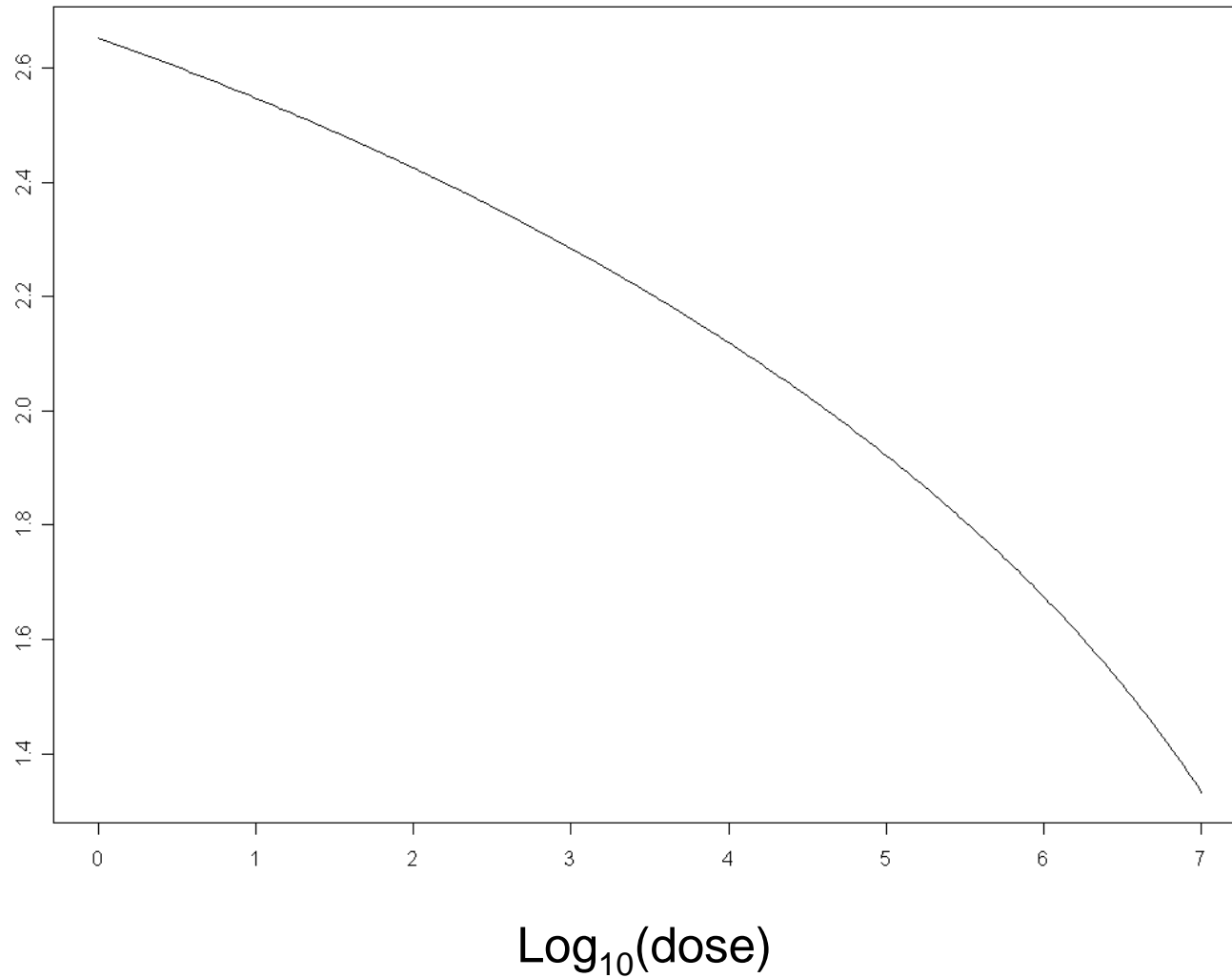
# Why Wilkening A1?

- One of three models (A1, A2, C) that Wilkening found fit Sverdlovsk data
- For Q fever, relationship is linear (as in A1), not curvilinear (as in A2)
- However, we could apply this method to any model of inhalational anthrax



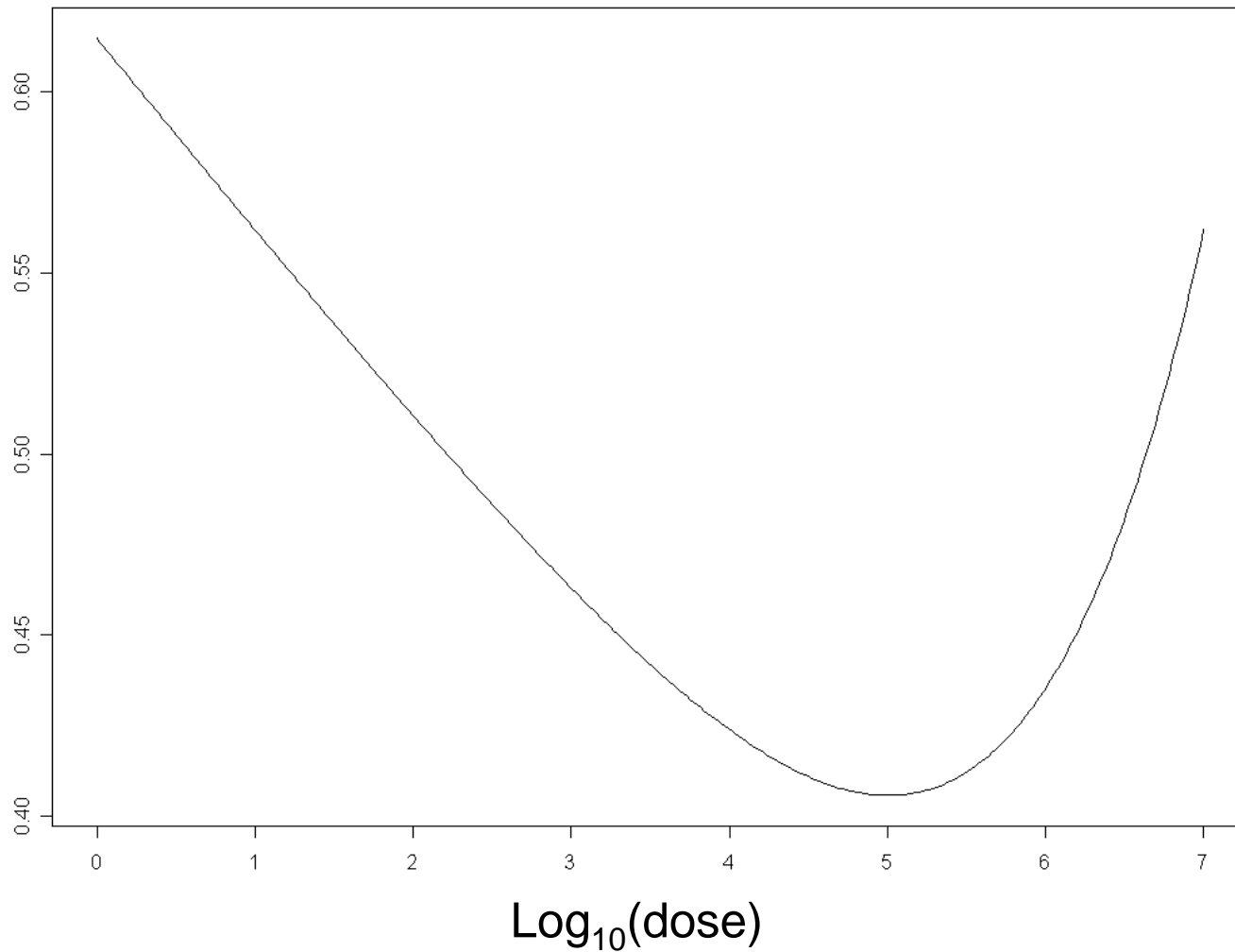
# Results

$\mu$  parameter for approximation  
(incubation +  
visit delay)



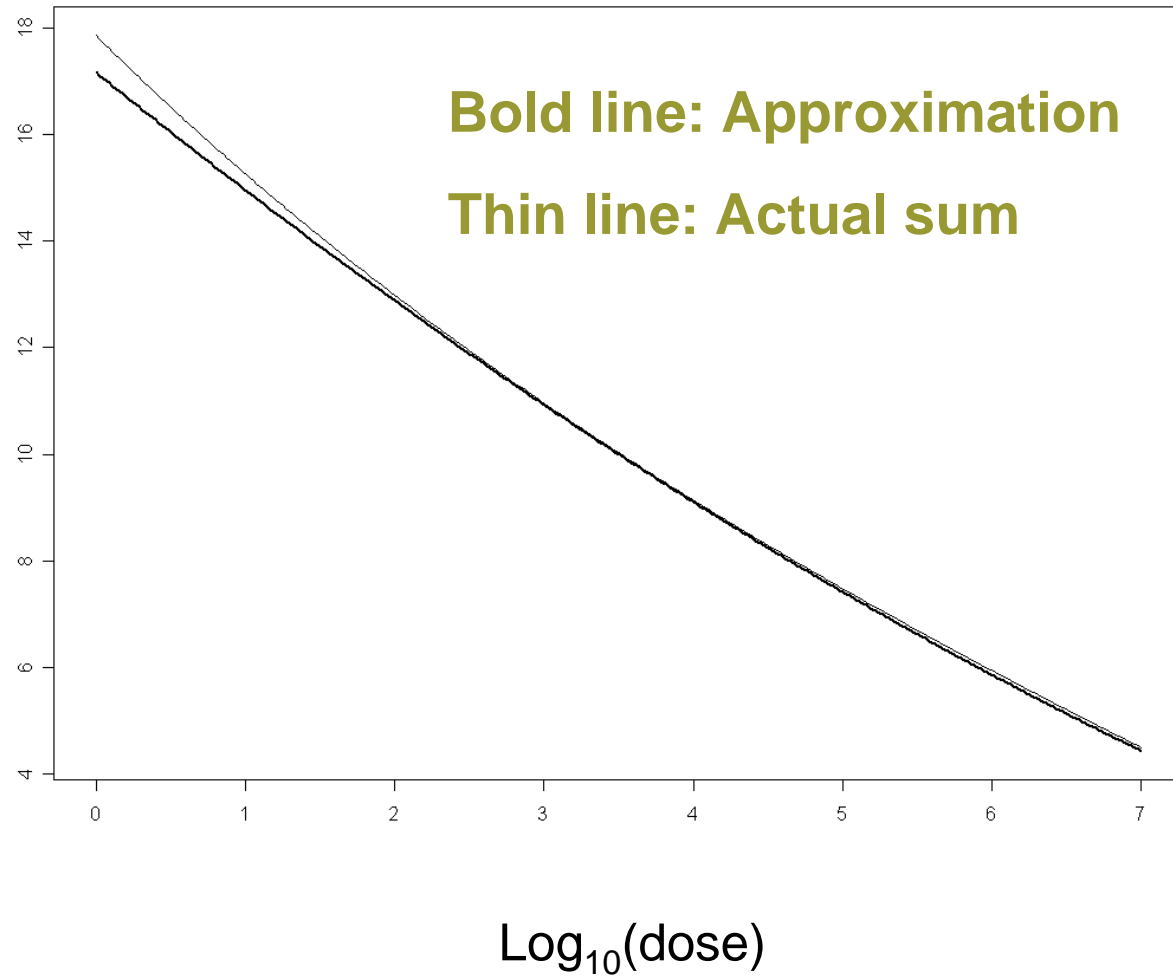
# Results (continued)

$\sigma$  parameter  
for  
approximation  
(incubation +  
visit delay)



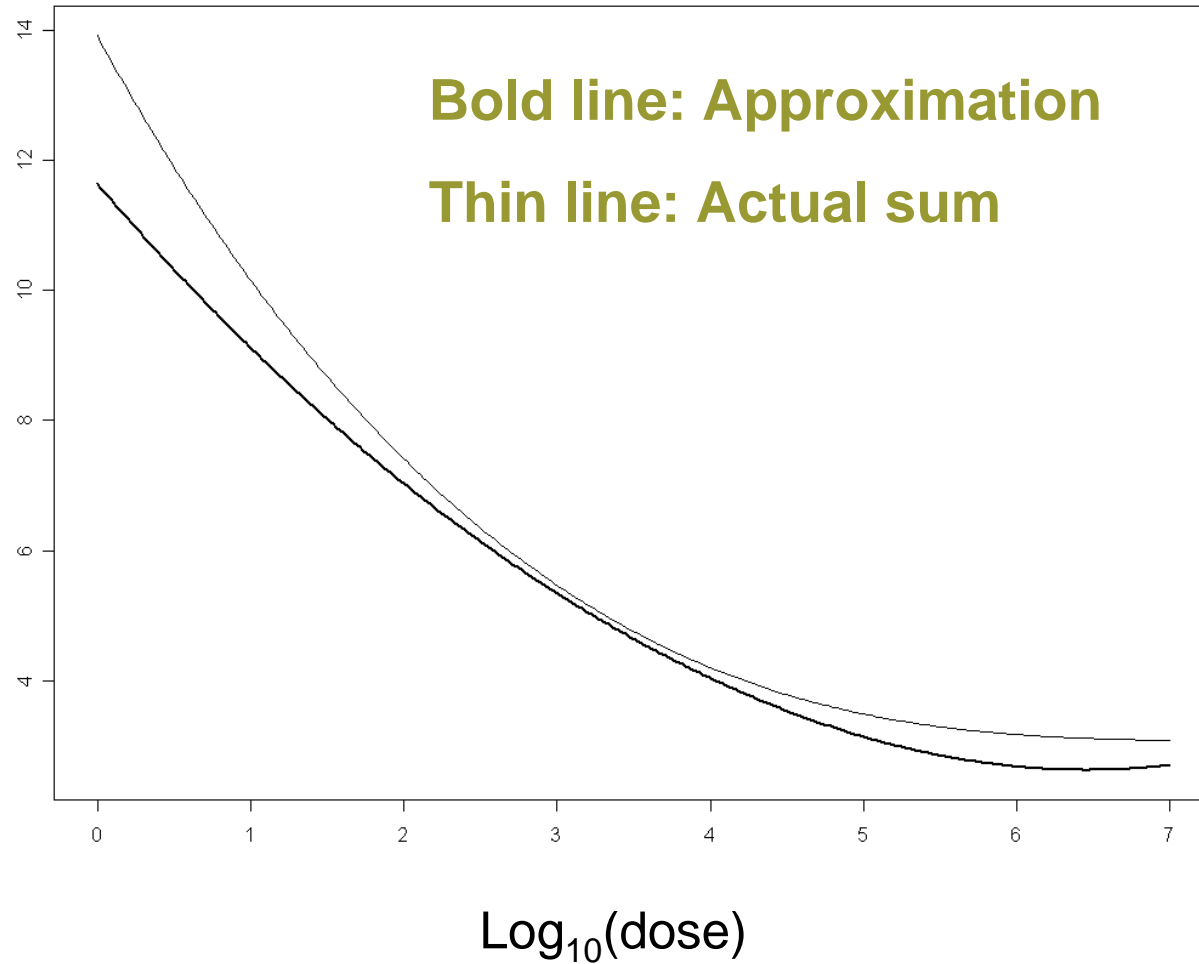
# Results (continued)

Mean of  
combined  
interval



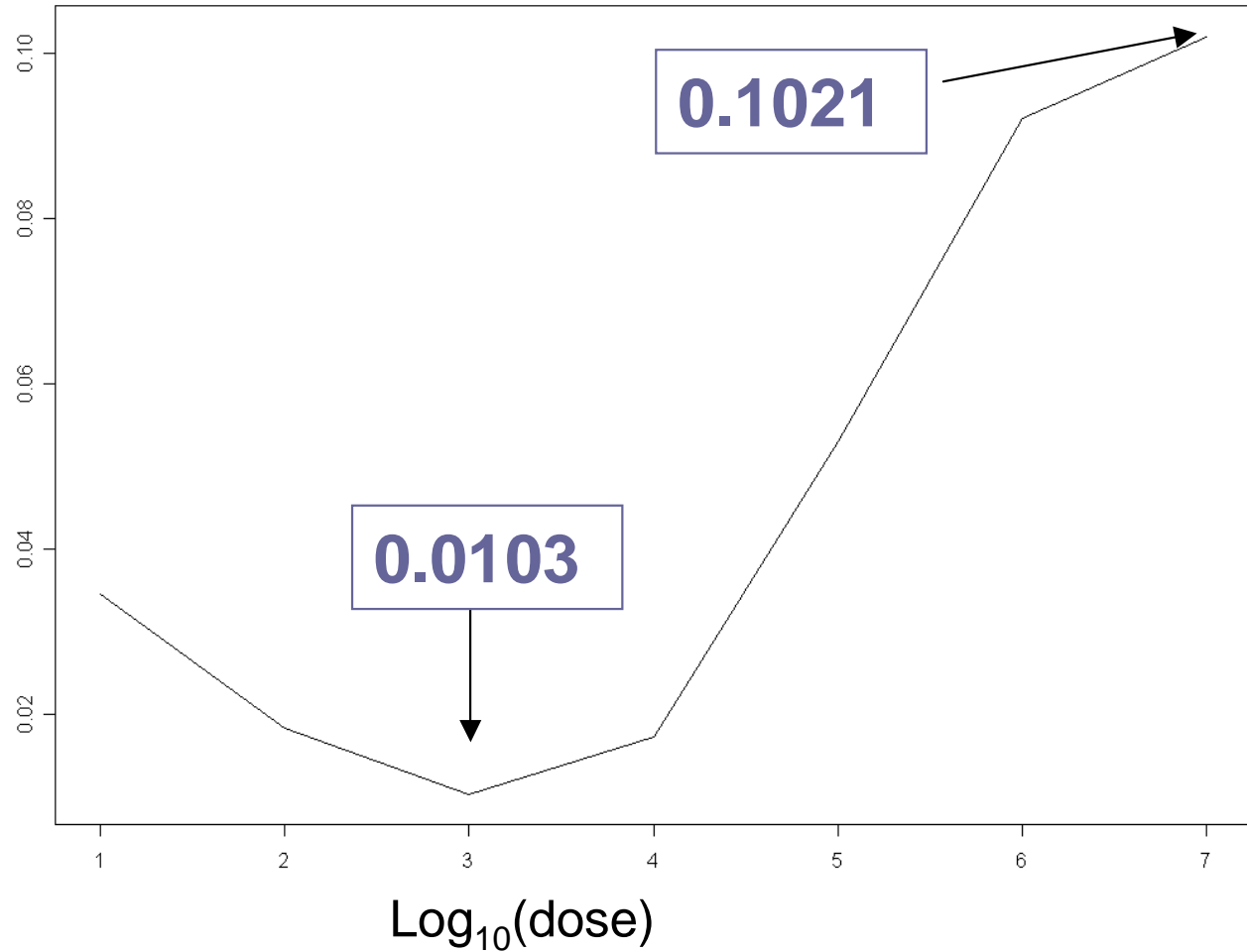
# Results (continued)

Std. dev. of  
combined  
interval



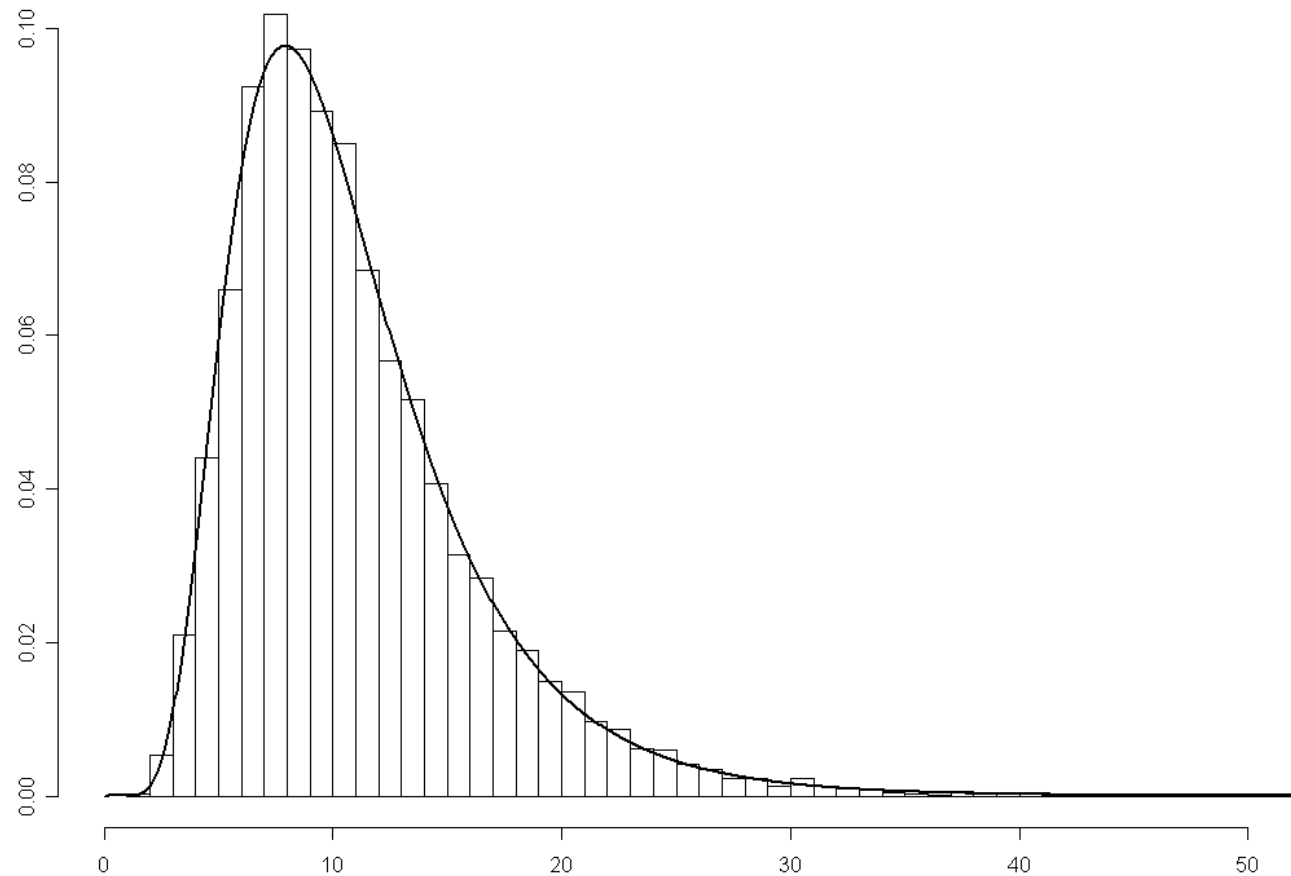
# Results (continued)

Hellinger distances between actual convolution and approximation



# Results (continued)

pdf of approx. vs  
histogram of  
draws from actual  
sum for dose =  
1000 spores\*

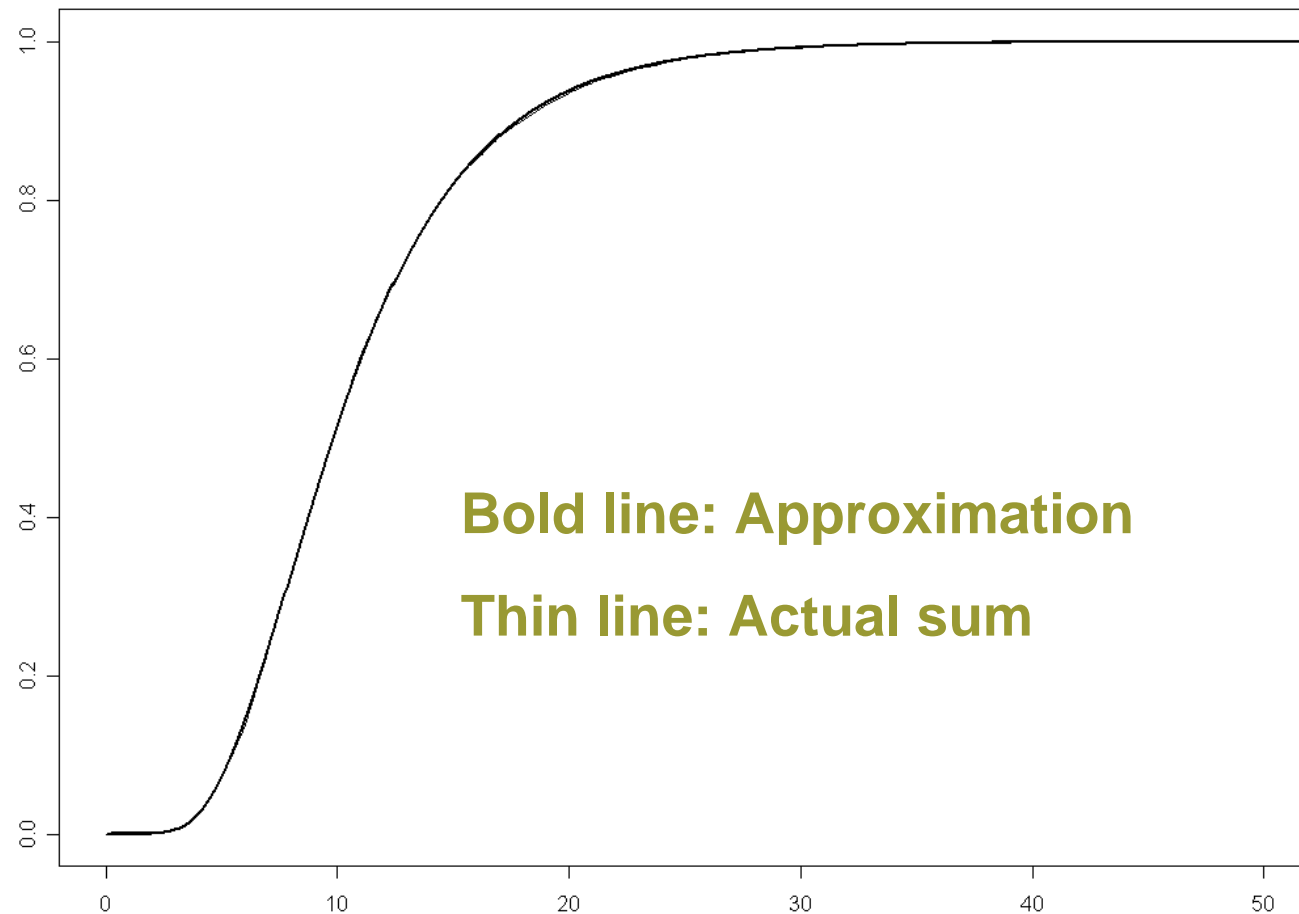


\* Shortest Hellinger distance



# Results (continued)

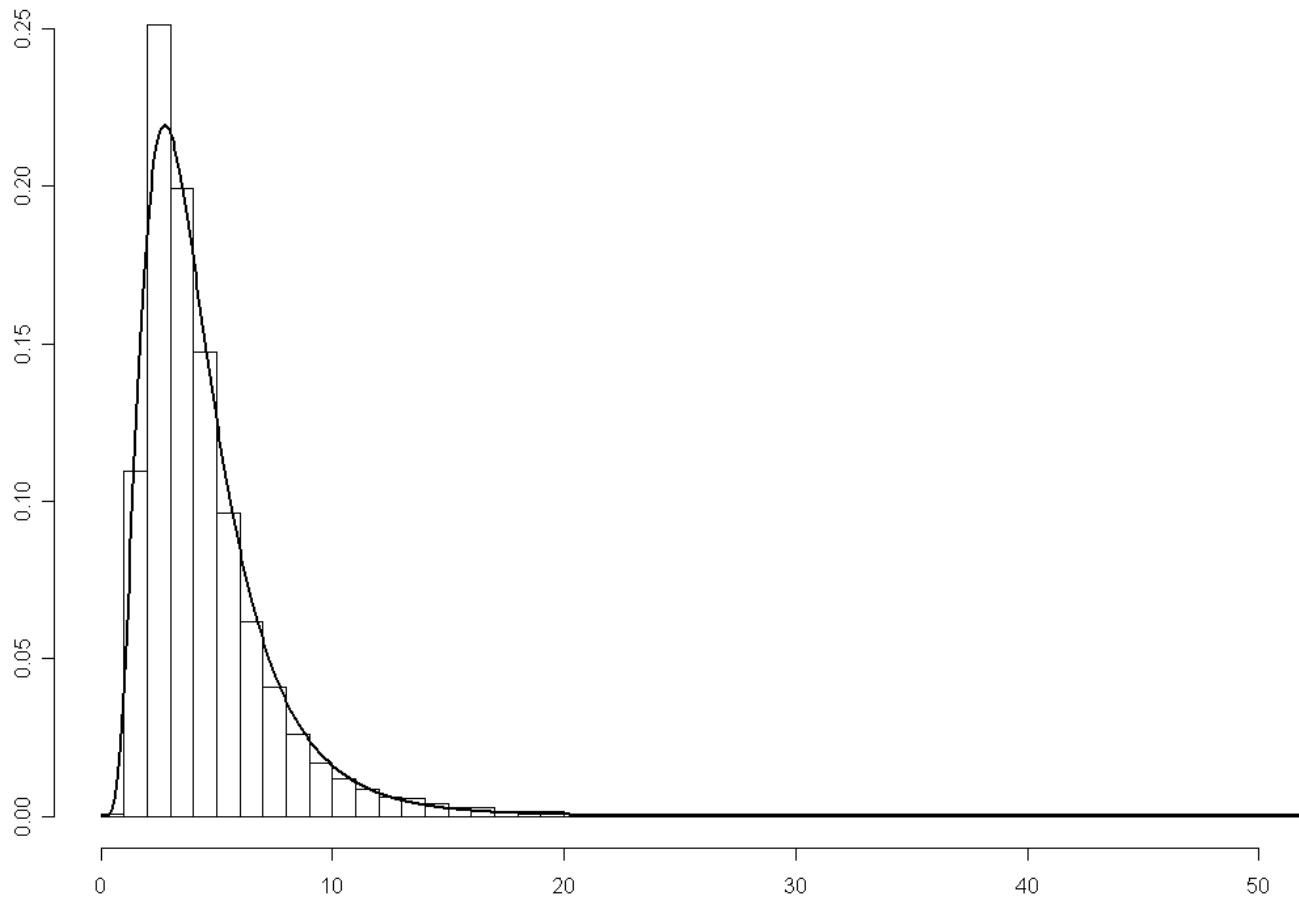
cdf of approx.  
for dose =  
1000 spores\*



\* Shortest Hellinger distance

# Results (continued)

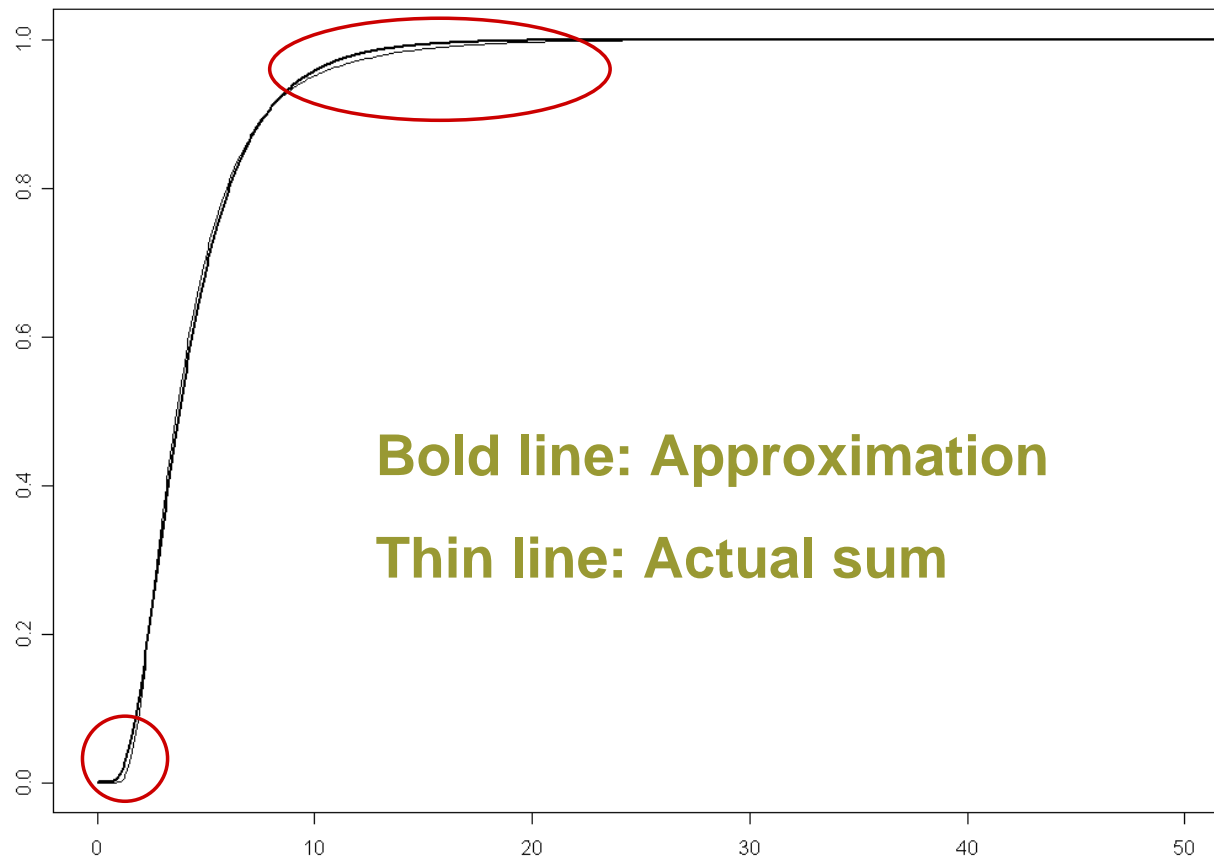
pdf of approx. vs  
histogram of  
draws from actual  
sum for dose =  
 $10^7$  spores\*



\* Longest Hellinger distance

# Results (continued)

cdf of approx.  
for dose =  
 $10^7$  spores\*



\* Longest Hellinger distance

---

## So What?

You just replaced a numerical convolution of two lognormals with a numerical solution to two equations for the  $\mu$  and  $\sigma$  parameters!

---

# Operational Use of the Approximation

- For the incubation period, we had simple formulas for computing  $\mu$  and  $\sigma$  of lognormal distribution
- Thus, we used spline interpolation to generate the following formulas for computing  $\mu$  and  $\sigma$  from  $\log_{10}(\text{dose spores})$ :

$$x = \log_{10}(\text{dose}), 1 \leq x < 7$$

$$w = \text{int}(x)$$

$$z = x - w$$

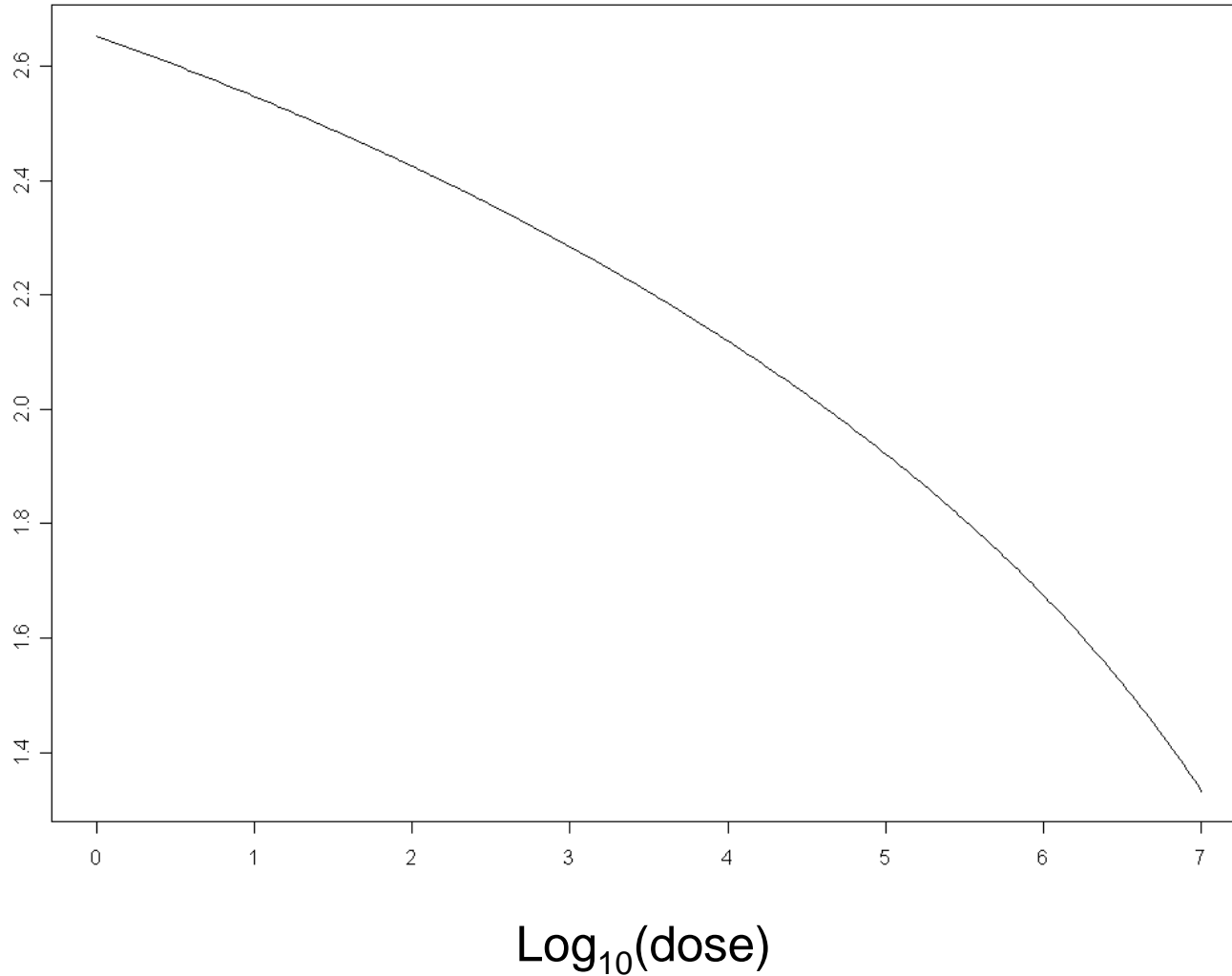
$$\mu = \beta_{1,w} \cdot z^3 + \beta_{2,w} \cdot (1-z)^3 + \beta_{3,w} \cdot z + \beta_{4,w} \cdot (1-z)$$

$$\sigma = \gamma_{1,w} \cdot z^3 + \gamma_{2,w} \cdot (1-z)^3 + \gamma_{3,w} \cdot z + \gamma_{4,w} \cdot (1-z)$$

For  $x < 1$  and  $x \geq 7$ , we just use  $x=1$  and  $x=6.99$ , respectively

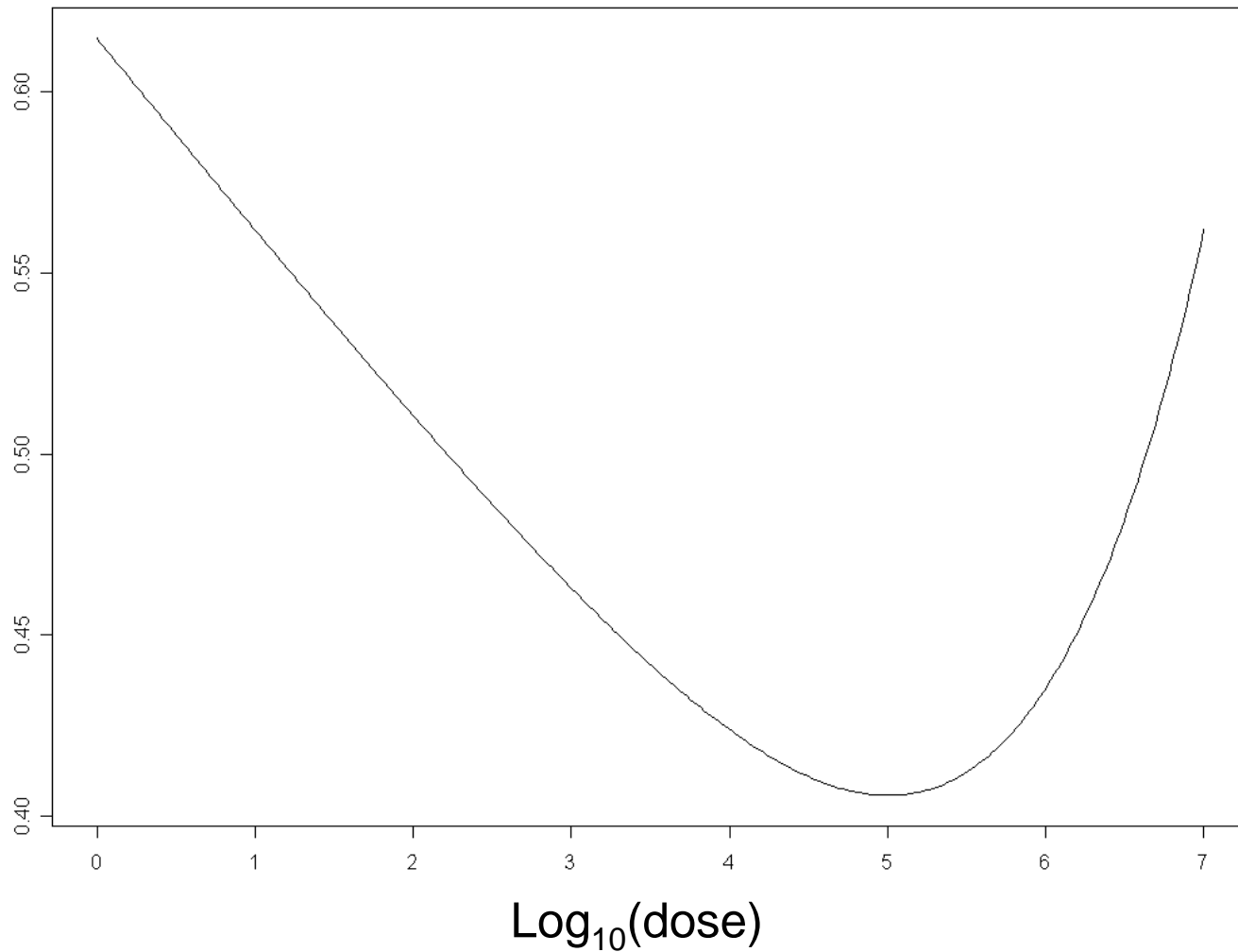
# Recall the Curve for $\mu$ Parameter

$\mu$  parameter for approximation  
(incubation +  
visit delay)



# ...and the Curve for the $\sigma$ Parameter

$\sigma$  parameter  
for  
approximation  
(incubation +  
visit delay)



# Operational Use: $\beta$ Coefficients

$w$	1	2	3	4	5	6
1	-0.0039	-0.0036	-0.0059	-0.0056	-0.0221	0.0000
2	0.0000	-0.0039	-0.0036	-0.0059	-0.0056	-0.0221
3	2.4292	2.2880	2.1253	1.9274	1.6961	1.3325
4	2.5474	2.4292	2.2880	2.1253	1.9274	1.6961

\* If you want more significant digits, email me

---



# Operational Use: $\gamma$ Coefficients

$w$	1	2	3	4	5	6
1	0.0007	0.0011	0.0037	0.0053	0.0229	0.0000
2	0.0000	0.0007	0.0011	0.0037	0.0053	0.0229
3	0.5100	0.4619	0.4202	0.4004	0.4124	0.5621
4	0.5620	0.5100	0.4619	0.4202	0.4004	0.4124

---

# Limitations

- Visit delay may vary with dose, although we're unaware of any data one way or the other
  - We assume that visit delay and incubation period are independent variables, given that an individual is infected
  - Visit delay may also vary with cultural, socioeconomic factors
-

---

# Conclusion

- We incorporated visit delay into Wilkening's A1 model of inhalational anthrax
  - The approximation of the sum of two lognormals was excellent for this application
  - We developed a simple spline formula for computing the parameters of the lognormal for the combined interval (incubation + visit delay) as a function of dose of inhaled spores
  - Operationally then, one can substitute the lognormal for the combined interval for the lognormal for incubation period, when desired and appropriate
-

---

# Acknowledgements

- Centers for Disease Control and Prevention grant 1 RO1 PH000026-01
  - National Science Foundation grant IIS-032558
-