

Some Properties of the Logarithmic and Exponential Functions and Their Applications in Pharmaceutical Studies

Jafar Soltani Farsani

Abstract: In this paper, a few properties of logarithmic and exponential functions will be reviewed by presenting problems arising from real studies. The properties are presented under 3 topics.

Introduction.

The logarithmic (In) and the exponential (exp) functions are two widely used mathematical functions in pharmaceutical studies. The In function is used to transform either the concentration data or PK parameters for various reasons, such as the log-down nature of concentration in the final phase or the need to create data with normal distribution for the PK parameters. After performing some statistical or mathematical procedures in the In-scale, the results are then back-transformed to the original scale using the exp functions.

In pharmaceutical studies, these two functions have properties that can result in some interesting and sometimes unexpected observations. In this paper, we present three topics. Under each topic, various properties of the aforementioned functions are reviewed. The first topic explains how, under special circumstances, the disproportionate scaling property of In and exp can create unusual values for the geometric and arithmetic mean of parameters. The second topic discusses 3 types of means that can be derived in pharmaceutical studies. Normally, pharmaceutical studies only present the geometric and arithmetic means. Using mathematical tools (and history!), we prove why the geometric mean is not a proper term to refer to the geometric means! We argue that a less popular term, i.e., the Geometric Least Square Mean (GLSM), is the right term for this type of mean. In the third topic, we present an observation which can be seen in many studies: the standard error of the In-transformed PK parameters is very



close to (sometimes identical with) the coefficient of variation of the original PK parameter. We will use mathematical tools such as the Taylor and McLaurin series expansions to prove why we make this observation.

Topic 1: Unusual values for the geometric and arithmetic means

Parameter (N_{A} / N_{B})	Geometric Means Arithmetic Means (CV %)				Ratio of Geometric Means	90% Confidence Interval	Intra- Subject CV (%)
	TRT A		TRT B		_		
AUCt (pg.h/mL) (19 /19)	2562.28 2709.00	(23.43)	4298.38 <mark>4812.23</mark>	(32.31)	59.61	52.59 - 67.57	18.11
AUCinf (pg.h/mL) (15 /16)	2860.01 3000.50	(23.85)	4386.56 <mark>4680.10</mark>	(33.63)	65.20	60.56 - 70.19	9.45

Figure 1 shows an unusual feature of arithmetic means for AUCt and AUCinf.

Figure 1- Unusual values for the arithmetic means for AUCt and AUCinf

The issue is that the arithmetic mean for AUCt is larger than that for AUCinf. This might seem unusual as AUCinf is an extension of AUCt meaning that for each individual, AUCinf is calculated by adding a positive value to AUCt. Therefore, AUCinf is always larger than AUCt.

Figure 2 provides the rationale behind this seemingly unexpected observation.



					AUCT	
			AUCT	AUCINF		CMAX
SUBJECT	PERIOD	GROUP	(pg.h/mL)	(pg.h/mL)	AUCINF	(pg/mL)
01	2	1	2680.13	2708.73	0.99	3898.94
02	2	1	7398.63	7500.53	0.99	9616.95
03	1	1	4759.49	4809.65	0.99	7316.69
04	1	1	2028.62	2087.91	0.97	2502.60
05	1	1	4614.82	4690.05	0.98	4085.53
06	2	1	5428.88	5510.12	0.99	6656.87
08	1	1	5130.29	5267.31	0.97	6542.73
09	1	1	6801.09			8643.87
10	2	1	6359.68			4821.23
11	2	1	4161.18	4200.75	0.99	5376.69
12	1	1	3697.44	3809.20	0.97	1941.60
13	2	1	7333.78	7483.01	0.98	10057.99
14	2	1	4592.45			5808.71
TMAX	LAMBDA	HALF-LIFE	LN (AUCT)	LN (AUCINF)	LN (CMAX)	AUC
(H)	(1/H)	(H)	(pg.h/mL)	(pg.h/mL)	(pg/mL)	(res%)
0.50	0.3640	1.90	7.8936	7.9042	8.2685	0.01
0.50	0.1495	4.64	8.9090	8.9227	9.1713	0.01
0.50	0.2229	3.11	8.4679	8.4784	8.8979	0.01
0.50	0.2425	2.86	7.6151	7.6439	7.8251	0.03
0.50	0.1578	4.39	8.4370	8.4532	8.3152	0.02
0.50	0.1455	4.76	8.5995	8.6143	8.8034	0.01
0.50	0.1062	6.53	8.5429	8.5693	8.7861	0.03
0.50			8.8248		9.0646	
0.50			8.7577		8.4808	
0.50	0.3329	2.08	8.3336	8.3430	8.5898	0.01
1.02	0.1209	5.73	8.2154	8.2452	7.5713	0.03
0.50	0.0963	7.20	8.9002	8.9204	9.2161	0.02
0.52			8.4322		8.6671	

Figure 2- Two subjects with missing AUCinf



Subjects 09 and 10 have AUCts not only much larger than the average AUCt, but also larger than the average AUCinf. However, since lambda could not be evaluated for these subjects (for reasons such as small R-squared value, missing time points in the final phase, etc.), AUCinf could not be calculated either. Therefore, they contribute to AUCt, but they add nothing to AUCinf. As a result, they shift the weight for arithmetic mean from AUCinf to AUCt.

Now, taking a closer look at Figure 1, another question arises: Why does the unusual observation for the arithmetic means doesn't apply to the geometric means? As seen in this figure, the geometric means have a natural order, with AUCinf having a larger geometric mean than AUCt. This observation might appear at odds with what was noted above for the arithmetic mean because of the following property of the logarithmic and exponential function.

Property 1:

exp and ln are strictly increasing functions.

This means that these two functions preserve orders. Hence, one might expect to observe the same order for the geometric and arithmetic means for AUCt and AUCinf. In other words, if the arithmetic mean for AUCt is larger than that of AUCinf, then so should be the geometric mean!

However, one should notice that in addition to property 1, the following property also plays a role in the discussed observation:

Property 2:

exp and ln do not keep differences proportionally.

More precisely, although the two functions preserve the orders, the magnitude of difference in the two scales (the original scale and the transformed scale) might be dramatically different. To visualize this fact, look at Figure 3.



					AUCT	
			AUCT	AUCINF		CMAX
SUBJECT	PERIOD	GROUP	(pg.h/mL)	(pg.h/mL)	AUCINF	(pg/mL)
01	2	1	2680.13	2708.73	0.99	3898.94
02	2	1	7398.63	7500.53	0.99	9616.95
03	1	1	4759.49	4809.65	0.99	7316.69
04	1	1	2028.62	2087.91	0.97	2502.60
05	1	1	4614.82	4690.05	0.98	4085.53
06	2	1	5428.88	5510.12	0.99	6656.87
08	1	1	5130.29	5267.31	0.97	6542.73
09	1	1	6801.09			8643.87
10	2	1	6359.68			4821.23
11	2	1	4161.18	4200.75	0.99	5376.69
12	1	1	3697.44	3809.20	0.97	1941.60
13	2	1	7333.78	7483.01	0.98	10057.99
14	2	1	4592.45		•	5808.71
TMAX	LAMBDA	HALF-LIFE	LN (AUCT)	LN (AUCINF)	LN (CMAX)	AUC
(H)	(1/H)	(H)	(pg.h/mL)	(pg.h/mL)	(pg/mL)	(rest)
0.50	0.3640	1.90	7.8936	7.9042	8.2685	0.01
0.50	0.1495	4.64	8.9090	8.9227	9.1713	0.01
0.50	0.2229	3.11	8.4679	8.4784	8.8979	0.01
0.50	0.2425	2.86	7.6151	7.6439	7.8251	0.03
0.50	0.1578	4.39	8.4370	8.4532	8.3152	0.02
0.50	0.1455	4.76	8.5995	8.6143	8.8034	0.01
0.50	0.1062	6.53	8.5429	8.5693	8.7861	0.03
0.50			8.8248		9.0646	
0.50			8.7577		8.4808	
0.50	0.3329	2.08	8.3336	8.3430	8.5898	0.01
1.02	0.1209	5.73	8.2154	8.2452	7.5713	0.03
0.50	0.0963	7.20	8.9002	8.9204	9.2161	0.02

Figure 3- Comparing the magnitude of AUCt in the original and log-transformed scales

.

8.6671

.

8.4322

.

.

0.52



Comparing AUCts for subjects 01, 09 and 10 we see the following order: 09, 10, 01. The exact same order applies to In(AUCt). However, AUCt for subjects 09 or 10 is much larger than AUCt for Subject 01 while a comparison between In(AUCt) values clearly shows that all 3 values are fairly close!

This property causes subjects 09 and 10 to lose ground after being log-transformed. They cannot offset the difference between mean values of ln(AUCt) and ln(AUCinf). As a result, the geometric mean (which is calculated as the back transformation of the LS-mean of log-transformed data) shows a natural order for AUCt and AUCinf.

Topic 2: On the geometric mean

The arithmetic mean and the geometric mean are two widely used types of means in statistics, particularly in pharmaceutical statistics. Figure 4 shows an example of such means reported for some BE study.

Parameter (N_{λ} / N_{B})	G Arith	Geometric Means Arithmetic Means (CV %)				90% Confidence Interval	Intra- Subject CV (%)
	TRT A		TRT B				
AUCt (ng.h/mL) (17 /17)	15692.49 16379.83	(27.75)	15268.67 <mark>16274.48</mark>	(31.45)	102.78	96.54 - 109.42	9.98
AUCinf (ng.h/mL) (17 /17)	16001.37 16679.81	(26.92)	15521.85 16531.24	(31.33)	103.09	97.03 - 109.52	9.65
Cmax (ng/mL) (17 /17)	<mark>1611.12</mark> 1634.46	(23.77)	1469.08 <mark>1562.13</mark>	(29.89)	109.67	98.76 - 121.78	16.77

Figure 4- Arithmetic and geometric means of PK parameters

Let's pick AUCt for TRT A as an example:



Arithmetic mean = 16379.83

Geometric mean = 15692.49

Exploring the outputs, we can extract a third kind of mean! Figure 5, shows the reported arithmetic mean for In(AUCt).

			AUCT							
Statistics	AUCT (ng.h/mL)	AUCINF (ng.h/mL)	AUCINF	CMAX (ng/mL)	TMAX (H)	LAMBDA (1/H)	HALF-LIFE (H)	LN(AUCT) (ng.h/mL)	LN(AUCINF) (ng.h/mL)	LN(CMAX) (ng/mL)
N	17	17	17	17	17	17	17	17	17	17
Max	26277.77	26852.95	0.9947	2319.34	8.00	0.1625	21.16	10.1765	10.1981	7.7490
Min	10375.22	11749.76	0.8806	1127.71	1.00	0.0328	4.27	9.2472	9.3716	7.0279
Median	15185.31	15486.86	0.9856	1476.76	2.50	0.1229	5.64	9.6281	9.6477	7.2976
Mean	16379.83	16679.81	0.9797	1634.46	3.36	0.1182	6.74	9.6695	9.6904	7.3733
Std Dev	4544.85	4489.38	0.0263	388.51	2.40	0.0330	3.90	0.2669	0.2546	0.2326
CV 8	27.75	26.92	2.68	23.77	71.37	27.94	57.79	2.76	2.63	3.16

Figure 5- The arithmetic mean for In(AUCt)

If we back transform this value using the exponential function, then we get a value which is close to both the arithmetic and geometric means but still distinct from them:

$$(9.6695) = 15827.43$$

Formulas for these three kinds of mean are presented below.

$$(equation 1) \qquad Arithmetic mean = \frac{1}{n} \sum_{i=1}^{n} a_i$$

$$(equation 2) \qquad Geometric mean = exp exp \left(\frac{1}{k} \left[\frac{1}{n_1} \sum_{i_1=1}^{n_1} \ln \ln (a_{i_1}) + \frac{1}{n_2} \sum_{i_2=1}^{n_2} \ln \ln (a_{i_2}) + \dots + \frac{1}{n_k} \sum_{i_k=1}^{n_k} \ln \ln (a_{i_k}) \right] \right)$$



(equation 3) The back transformation of the Arithmetic mean of the ln – transformed data
=
$$exp exp \left(\frac{1}{n} \sum_{i=1}^{n} ln ln (a_i)\right)$$

At this point, let's shift the main theme from statistics to mathematics. The original mathematical formula for the geometric mean is shown below:

Geometric mean =
$$\sqrt[n]{a_1 \times a_2 \times ... \times a_n} = \left(\prod_{i=1}^n a_i\right)^{\frac{1}{n}}$$

One might wonder why this formula is called the geometric mean. The notion of geometric mean originates from triangular geometry.

Figure 6 depicts a right triangle,



Figure 6- A right triangle with height h

Historically, the following relationship has been known for this triangle

$$h^2 = xy$$



$$\Rightarrow h = \sqrt{xy} = (xy)^{\frac{1}{2}}$$

This is the formula for the geometric mean for n=2. Although this geometric feature applies only to n=2, the formula is generally called the geometric mean for any value of n. It is worth noting that there is a second geometric interpretation for n=2: The geometric mean of two numbers is the length of each side of a square, the area of which is the same as the area of a rectangle with its length and width being equal to the two given numbers. Also, for n=3 there is a geometric interpretation: The geometric mean of 3 numbers is the length of each side the volume of which is equal to the volume of a cube with the given 3 numbers as its width, length, and height. There are no geometric interpretations for values of n larger than 3.

The formula provided above seems still unfamiliar to what we know as the geometric mean. However, we can use the following 3 properties of the In function to convert the formula for geometric mean to something we already know.

Property 3:

exp exp (ln ln (a)) = a, for any a > 0.

(exp exp (a)) = a, for any value of a.

Property 4:

 $(a^b) = b \times ln(a)$, for any a > 0 and any value of b.

Property 5:

 $\ln \ln (a_1 \times a_2 \times ... \times a_n) = \ln \ln (a_1) + \ln \ln (a_2) + \dots + \ln \ln (a_n).$

Now, applying these properties consecutively, we can write



$$\sqrt[n]{a_1 \times a_2 \times \dots \times a_n} = \left(\prod_{i=1}^n a_i \right)^{\frac{1}{n}}$$
$$= exp \ exp \ \left(ln \ ln \left(\left(\prod_{i=1}^n a_i \right)^{\frac{1}{n}} \right) \right)$$
$$= exp \ exp \ \left(\frac{1}{n} \ ln \ ln \left(\prod_{i=1}^n a_i \right)^{\frac{1}{n}} \right) \right)$$
$$= exp \ exp \ \left(\frac{1}{n} \sum_{i=1}^n ln \ ln \ (a_i) \right)$$

Therefore, we see that, mathematically, the geometric mean is equal to the back-transformation of the arithmetic mean of the Intransformed data (equation 3). What we referred to as the geometric mean of data before (equation 2) has nothing to do with the recent formula! Actually, it is important to note that what we typically call the geometric mean is not a geometric mean! This is just an abuse of terminology! This is why some references refer to (equation 2) as the Geometric Least Square Mean (GLSM). The reason is that the formula is derived from the least square mean given by a GLM or Mixed model.

Still many references use the term Geometric mean to refer to equation 2. The reason is that equation 3 has no practical use in statistics. Basically, we transform our data using the In function because we hope the In-transformed data would be normally distributed. This will allow us to use ANOVA and GLM/Mixed models. Recall that normality of data is the basic requirement of these methods. On the other hand, these methods always return the LS mean (the Lease Square mean) as the point estimation for average. We arrive at the conclusion that In-transformation and the GLSM formula (equation 2) come as a package, and we cannot



separate them. This means that equation 3 doesn't provide any insight or valuable information. Hence, its corresponding term (the geometric mean) is sometimes taken away and assigned to equation 2!

Topic 3: An immediate estimation for the Coefficient of Variation

Reviewing the results of studies analyzed using a GLM model, one can identify the fact that typically the RMSE reported by the ANOVA for the In-transformed data and the CV for the original PK parameters are very close and sometimes identical.

Figures 7-9 display this fact for 3 parameters, AUCt, AUCinf and Cmax, from the same study. If we round off RMSE to 4 decimal places and then multiply it by 100 (to change the unit to %), then we get values which are identical to CV for AUCt and AUCinf, and close to CV for Cmax.



The GLM Procedure

Dependent Variable: LNAUCT

		Sum of			
Source	DF	Squares	Mean Square	F Value	Pr > F
Model	24	1.99751912	0.08322996	57.05	<.0001
Error	21	0.03063825	0.00145896		
Corrected Total	45	2.02815736			

1	R-Square	Coeff Var	Root MSE	LNAUCT Mean
(0.984894	0.876844	0.038196	4.356121

Source	DF	Type III SS	Mean Square	F Value	Pr > F
SEQ	1	0.19378029	0.19378029	132.82	<.0001
SUBJECT (SEQ)	21	1.79970287	0.08570014	58.74	<.0001
PERIOD	1	0.00190368	0.00190368	1.30	0.2662
TREATMENT	1	0.00195682	0.00195682	1.34	0.2598

Tests of Hypotheses Using the Type III MS for SUBJECT(SEQ) as an Error Term

Source	DF	Type III SS	Mean Square	F Value	Pr > F
SEQ	1	0.19378029	0.19378029	2.26	0.1475

Figure 7- RMSE (In(AUCt)) and CV(AUCt) are identical

Figure 8- RMSE (In(AUCinf)) and CV(AUCinf) are identical



The GLM Procedure

Dependent Variable: LNAUCT

Source		DF	Sum Squa:	of res	Mean Squar	e F	Value	Pr > F
Model		24	1.99751	912	0.0832299	6	57.05	<.0001
Error		21	0.03063	825	0.0014589	6		
Corrected Total		45	2.02815	736				
	R-Square	Coeff	Var	Root M	SE LNAUC	T Mean		

0.984894	0.876844	0.038196	4.356121

Parameter (N _A /N _B)	Arit	Geometric M thmetic Mear	feans ns (CV දි)		Ratio of Geometric Means	90% Confidence Interval	Intra- Subject CV (%)	
	TRT A		TRT B					
AUCt	78.69		77.67		101.31	99.37 - 103.30	3.82	
(µg.h/mL) (23 /23)	80.14	(20.79)	79.24	(22.53)				
AUCinf	83.61		82.73		101.06	99.05 - 103.12	3.96	
(µg.h/mL) (23 /23)	85.06	(20.15)	84.30	(21.93)				
Cmax	18.04		17.39		103.74	97.56 - 110.31	12.14	
(µg/mL) (23 /23)	18.36	(21.56)	17.73	(20.65)				

Figure 9- RMSE (In(Cmax)) and CV(Cmax) are very close



Below, the mathematical justification for this observation is provided. The approach is based on a well-known mathematical tool called series expansion and polynomials.

Let f be a function that meets a couple of mathematical requirements. Without getting into those mathematical details, keep in mind that both the logarithmic and exponential functions meet the requirements. The Taylor series expansion is defined as follows:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n$$
$$= f(a) + f'(a)(x-a) + \frac{f''(a)}{2} (x-a)^2 + \frac{f^{(3)}(a)}{6} (x-a)^3 + \cdots$$

In this formula, a is a fixed point, which we will refer to as the center of expansion. x is any arbitrary point in the range of f. $f^{(n)}(a)$ denotes the n-th derivation of f at a. The formula above says, instead of calculating f(x) directly, we can use the formula on the right-hand side. The latter could be simpler to apply, as it is a series of polynomials.

One issue regarding this formula is that it provides an infinite series, which is impractical. Therefore, the notion of the K-th order Taylor polynomial has been introduced as shown below.

$$f(x) \approx \sum_{n=0}^{k} \frac{f^{(n)}(a)}{n!} (x-a)^{n}$$

In this formula k represents any arbitrary integer (i.e. k=1, 2, ...). Since we are throwing away parts of the original Taylor series, we do not get a 100% accurate equality, but we get an estimation instead. The accuracy of this estimation depends on two factors:

- 1. How large k is: The larger the k, the more accurate the formula.
- 2. How close *x* is to *a*: The closer the x is to a, the more accurate the formula.



Also note that there is an interaction between 1 and 2: If we choose a larger k, then we can pick any value of x possibly farther from the center of expansion (a) and still get an accurate estimation. However, if we only want to use the formula to estimate f(x) for some x in the neighborhood of a, then a small value of k suffices.

The Maclaurin Series expansion is a special case of the Taylor series expansion, where a = 0.

Property 6:

- Maclaurin series expansion for the exponential function:

$$exp \ exp \ (x) = \sum_{n=0}^{\infty} \frac{1}{n!} x^n$$

Property 7:

- A couple of Maclaurin polynomials for the exponential function:
- The 1st-order Maclaurin polynomial:

 $exp \ exp \ (x) \approx 1 + x.$

• The 2nd-order Maclaurin polynomial:

$$exp \ exp \ (x) \ \approx 1 + x + \frac{1}{2}x^2.$$

• The 3rd-order Maclaurin polynomial:

$$(x) \approx 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3.$$

Now we have all the mathematical tools to prove the relationship specified in the beginning of this topic. Let's pick the 1st-order Maclaurin series for exp and let $x = \sigma^2$, the variance of the In-transformed PK parameter, then



 $exp exp (\sigma^{2}) \approx 1 + \sigma^{2}$ $\Rightarrow \sigma^{2} \approx exp exp (\sigma^{2}) - 1$ $\Rightarrow \sigma \approx \sqrt{exp exp (\sigma^{2}) - 1}$ $\Rightarrow SD(lnPK) \approx CV(PK)$

SD(InPK) stands for the standard deviation of the In-transformed PK parameter. Notice that RMSE of the In-transformed PK parameter is an estimation for SD(InPK). This proves the observation under discussion!

One question that might arise at this point is that, with picking the very basic McLaurin polynomial (i.e. the 1-st order polynomial), we might not get an accurate estimation. However, the point is that since the variance of the In-transformed PK parameter is typically very small (close to 0), and 0 is nothing but the center of expansion for the McLaurin polynomials, we get a good estimation even with the 1st -order polynomial.

Let us revisit the example provided in the beginning of this topic,

SD(lnAUCt) = 0.038196		, $CV(AUCt) = \%3.82 \ (= .0382)$
SD(lnAUCinf) = 0.039573	,	CV(AUCinf) = %3.96 (= .0396)
SD(lnCmax) = 0.120978	,	$CV(Cmax) = \%12.14 \ (= .1214)$

Now, we can say why SD(InPK) and CV(PK) for AUCt and AUCinf perfectly match while for Cmax we don't have identical values (although they are fairly close). The reason is that SD(InPK) for the first two parameters are closer to 0 (the center of expansion). Consequently, the estimations for these parameters are more accurate than that for Cmax.

We can delineate intervals in terms of accuracy of this estimation as shown below:

1. SD(lnPK) in (0, .0736]: $SD(lnPK) \approx CV(PK)$ to 4 decimal places



- 2. SD(lnPK) in (.0736, .1220]: $SD(lnPK) \approx CV(PK)$ to 3 decimal places,
- 3. SD(lnPK) in (.1220, .2888]: $SD(lnPK) \approx CV(PK)$ to 2 decimal places,
- 4. SD(lnPK) in (.2888, .5140]: $SD(lnPK) \approx CV(PK)$ to 1 decimal place.

Please note that decimal places are assumed to be represented after rounding off.

Why Choose BioPharma Services?

At Biopharma Services, we use the most appropriate tools and methods to analyze the collected data. We provide scientific justifications for any observation, either expected or unexpected, based on the statistical and mathematical tools. Our Stats team is committed to address all the questions and concerns from our valued sponsors using expert topics like <u>CNS Drugs</u>, <u>early phase</u> clinical trials, oncology drug development explained in a simple and relatable language.

BioPharma Services, Inc., a Think Research Corporation and clinical trial services company, is a full-service Contract Clinical Research Organization (CRO) based in Toronto, Canada, specializing in Phase 1 clinical trials 1/2a and Bioequivalence clinical trials for international pharmaceutical companies worldwide. BioPharma has clinical facilities both in the USA and Canada with access to healthy volunteers and special populations.