Effect of Activity Patterns on Radon Exposure

Willard E. Hobbs, Ph.D.
Radon Reduction & Research
Santa Barbara, CA 93105

ABSTRACT

An important goal of those concerned with air pollution is to estimate the time-weighted average exposure to various indoor air pollutants. Various studies have measured the concentration of pollutant for indoor environments, but without regard to the presence of humans. Other studies have used diaries and other records to evaluate the activity patterns relative to various environments, indoor and outdoor. In this presentation I will show how the results of these studies can be combined analytically to provide estimates of the distribution of human exposures to radon. This is a big step beyond the earlier approaches, which have used random sampling to computationally estimate the exposure distribution (Ref. 1). As an example, I will use residential radon concentrations and activity patterns associated with people in the State of California. For both populations, homes and people, there have been detailed studies made using samples of size larger than 3,000. The relatively large sample size was needed to consider various strata. For homes, they considered such things as housing type, climate, and geologic province. For people, they considered such things age, gender and income level. The salient points of these studies will be reviewed. In both cases state officials are satisfied that the obtained distributions accurately reflect the populations. I focus on the residential environment because it is the environment where we spend the most time and the one where people receive their largest dose of natural radiation (from radon, of course). The procedure presented is to evaluate the exposure distribution by averaging over all home concentration levels and all human activity patterns. It reduces to a type of convolution of these two distributions. The home radon concentration levels are well approximated as a lognormal distribution and the at-home fraction of a day may be approximated as a histogram. Corrections are made for the concentrations because people spend the nighttime hours at home when radon levels are twice their day-average recorded levels. With these assumptions, an analytic formula for the daily human exposure distribution per unit radon (pCi-day/liter) may be analytically calculated. It is a single peak distribution which is broader that a lognormal distribution. Albeit, the arithmetic average for California is a modest 0.9 pCi/L with less than one percent of population predicted to have a time-weighted exposure greater than 4.0 pCi/L. Because the procedure is analytic, it may be evaluated using ubiquitous spreadsheet software requiring minimal computer resources. It would be interesting to compare the recent updates to the California activity patterns with those found in Iowa (reported at last year's AARST meeting). Nevertheless, it is not believed that activity patterns vary significantly from state to state, and therefore, Californian patterns can be used to estimate human exposure for radon in other states.

1. M. D. Koontz, W. C. Evans, and C.R. Wilkes, "Development of a Model for Assessing Indoor Exposure to Air Pollutants," GOEMET Technologies, Inc., Report A933-157, California Air Resources Board, 2020 L St, Sacramento, CA, January 1998.

INTRODUCTION

An important goal of those concerned with air pollution is to estimate the time-weighted average exposure to various indoor air pollutants. Numerous studies have measured the concentration of pollutants for indoor environments, but usually without regard to the presence of humans. Other studies have used diaries and records to evaluate the activity patterns and access the time in various environments both indoor and outdoor. In this paper I will show how the results of these studies can be combined analytically to provide accurate estimates of human exposures. I will use activity patterns and residential radon concentrations associated with people and houses in the State of California.

I consider the residential environment because it is the one where we spend the most time and the one where people receive their largest dose of radon radiation. The average Californian spends approximately 17.2 hours per day in a residence. Because of the design and use of homes, of all building types, the exposure to radon is highest in residences. In homes radon concentrations average an order of magnitude greater than outdoor levels. High nighttime levels driven by the stack effect dominate the high residential levels. A factor relating to high residential radon levels is that ventilation systems in homes are not regulated. Air exchange often occurs inadvertently (by default) allowing unhealthy conditions to develop and linger. Air handling units in commercial and public building are regulated, maintained, and used. This usually keeps radon concentrations in those buildings at negligible levels.

The procedure presented here is to convolute the indoor radon concentration distribution with the activity patterns that show the distribution of time spent in residences. The residential concentration distribution uses data generated by the California Department of Health Services and time-dependent data that I gathered. Like many indoor and outdoor pollutants radon concentrations are accurately modeled with a lognormal distribution function. The mean radon concentration is increased by a factor to account for the higher nighttime levels when the occupants are present. We begin the next section with a brief discussion of the data and its approximation. The activity profiles are the focus of this paper. We make three approximations of the distribution of time spent at home by the people of California. Each divides the day into different subsets of portions of day spent at home. The division doesn't make a significant difference, as the resulting distribution of exposures is essentially the same for each. In the conclusion the way the exposures vary with the details of the assumed input functions and other implications of this analysis are discussed.

CALIFORNIA DATA

Residential Radon Concentrations There have been two major statewide surveys of residential radon concentrations in California. The first by Liu, *et al.*[1] involved 310 year-long alpha-track measurements and the second by Quinton[2] involved about 3000 charcoal canisters in summer and autumn short-term screening measurements. In both of these data sets 1.9% of the measurements were above 4.0 pCi/L. The reason for the close agreement is the time-of-year for the screening measurements, occurring during the hottest portion of the year when the stack effect is weakest.

Hobbs and Meada[3] have studied the data sets and found that the best parameters are a geometric mean of 0.65 pCi/L and a geometric standard deviation factor of 2.4. Their method for analyzing the data does not include measurements with magnitude less than 1.0 pCi/L and is designed to predict the proportion of homes with elevated radon concentrations. For example, the above parameters give a fraction of homes with radon levels above 4.0 pCi/L as 1.9%, agreeing with the observations.

In this paper the geometric mean found by Hobbs and Meada[3] is increased to 1.08 pCi/L to account for the higher nighttime levels. Nero, *et al.*[4] have shown that radon levels in houses follow a lognormal distribution. A simple approach to such a distribution is to note that the logarithms of the measurements follow a standard normal (Gaussian) distribution. The geometric standard deviation is therefore a factor within which approximately 68% of the distribution lies. (For the assumed California distribution this would be between 1.08/2.4 = 0.45 pCi/L and 0.65x2.4 = 2.59 pCi/L or approximately 0.5 – 2.6 pCi/L.)

This paper considers situations where the population of homes is well characterized by a single lognormal distribution. The Hobbs and Meada[3] analysis found that regions with large geometric standard deviations were often associated with radon "hot spots." These are situations where the distribution of radon concentrations is characterized by a sum of two lognormal distributions, one with a significantly larger geometric mean. Here, I assume that this is not the case.

Activity Patterns The California Air Resources Board has sponsored two surveys whereby 24-hour time diaries were administered to a modified random sample of California residents. The first by Jenkins, *et al.*[5] involved a target population of adults and adolescents and provided 1,762 location/activity profiles. The second by Phillips, *et al.*[6] involved a target population of children and provided an additional 1,200 location/activity profiles, for a total of 2,962 profiles. The two surveys are complimentary in that virtually identical schemes were used for recording both location/activity information and selected characteristics of the participants.

For this research I accessed the California activity profiles using the computer code CPIEM (California Population Indoor-Air Exposure Model) developed by Koontz, *et al.*[7] The code itself was not run; only the 24-hour activity input file was used. This gave the time spent in the participant's residence as well as other persons' residences. Other location codes are included as well as activity levels but these were not used. Various demographic strata have been analyzed and the data are believed to be an accurate representation of California activities.

TIME IN RESIDENCE AND RADON EXPOSURE

To calculate the average exposure conceptually I make a choice of a random Californian. The radon in his house will be randomly chosen from the distribution of indoor radon levels and the fraction of time he spends in the house will be randomly chosen from the distribution of activity profiles. I assume that these two factors are independent. The time in residences is represented as a fraction of a day f between 0 and 1. The distribution must include the two limits because there are individuals (relatively few) who do not spend any time in a residence $f=0$ and others (relatively many) who do not leave home $f=1$.

The distribution of time-average radon exposure for the individuals who do not leave home is simply the distribution of radon concentrations in California homes. Since we are calculating residential exposure, people who do not enter a residences, do not receive any exposure. Consider the subset of the population that spends exactly half their day in a residence. For this fraction of people to receive a time-average exposure of 2 pCi/L, they would need to be in homes with an average radon concentration of 4 pCi/L. Recall that the radon geometric mean has been increased because the half a day that individuals spend in residences usually includes the nighttime when radon levels are higher. The details of how the distributions are convolution are included in the Appendix.

Delta-Function Approximation Figure 1 shows an approximation to the distribution of day fraction from the California surveys. It is represented as a bar chart where the height of each bar is the number of records reporting a given percent plus or minus one-half percent. The top bar has the largest value; 181 respondents report spending more that 0.995 of a day at home. (A closer examination shows that 164 spent all day at home while 17 reported spending less than 0.005 day = 7 min 12 sec not at home.)

Assuming that the number of people at each percentage spent exactly that fraction of a day in a residence, the exposure to the population can be calculated. The lognormal distribution for the radon concentration x in California homes is

$$c(x,\mu,\sigma)=\frac{\exp[-(\ln x-\mu)^2/2\sigma^2]}{x\sigma\sqrt{2\pi}} \qquad (1)$$

Where μ and σ are the natural logarithms of the geometric mean and geometric standard deviation. If f_n is the fraction of the population that spends fraction y_n of the day in residences, the distribution function is $a_n(y)=f_n\,\delta(y-y_n)$ and the corresponding exposure $e_n(x)$ is

$$e_n(x)=\int_0^1 a_n(y)c\left(\frac{x}{y}\right)\frac{dy}{y}=\frac{f_n c(x/y_n,\mu,\sigma)}{y_n}=f_n c(x,y_n\mu,\sigma) \qquad (2)$$

N.B., the expression $y_n\mu$ is the log of the product of day-fraction y_n and the geometric mean e^{μ} (see footnote 8). Summing over all the fractions for each percent of day, the total exposure distribution is

$$e(x)=\sum_{i=1}^{100} f_i c(x,y_i\mu,\sigma) \qquad (3)$$

This is a sequence of lognormal functions with decreasing means and amplitudes equal to the relative fractions of the population. The resulting exposure distribution and the original home concentration distributions are shown in Figure 2.

The distribution function for the fraction of a day spent in a resident by a Californian is the Green's function which, when convoluted with the home radon concentration distribution, gives the exposure distribution. In the above analysis the Green's function is approximated as a sum of 101 delta functions. One approach at this point would be to find a smooth function that could then be analytically convoluted and the exposure thereby evaluated.

Histogram Approximation I have chosen to approximate the day-fraction function in even simpler terms. Figure 3 shows a histogram for the day-fractions divided into 10 bins, each with a width 10%, plus a bottom bin for 0% and a top bin for 100%. [The 164 people with activity profiles showing that they do not leave home are in this top category. They are not a bin but are represented by a delta function $F_{100\%}\delta(1-y)$.] This histogram expression may be evaluated using the lognormal cumulative distribution function

$$Ncdf(a,\mu,\sigma)=\int_0^a \frac{\exp[-(\ln x-\mu)^2/2\sigma^2]}{x\sigma\sqrt{2\pi}}dx \qquad (4)$$

Where the symbols have the same meaning as before. The result for the exposure distribution is

$$e(x) = F_{100\%}c(x,\mu,\sigma) + \sum_{i=1}^{10} n_i F_i \left[Ncdf\left(10x/(i-1),\mu,\sigma\right) - Ncdf\left(10x/i,\mu,\sigma\right) \right]$$

or (5)

$$e(x) = F_{100\%}c(x,\mu,\sigma) + \sum_{i=2}^{10} (n_{i-1}F_{i-1} - n_i F_i) Ncdf(x, i\mu/10, \sigma) \; + \; n_1 F_1$$

The n_i factors normalize the cumulative difference functions and F_i are the proportions of the population in the 10 bins. This exposure and the residential concentration distribution are also shown in Figure 2. It is very close to the exposures found using the delta functions.

Step-Function Approximation An even simpler approximation for the day-fraction function would be a single step function $S(y,a,b)$. S is equal to 1 when y is between a and b.

$$a(y) = n_s \; S(y, 0.4, 1.0) \qquad (6)$$

Now the exposure is evaluated

$$e(x) = n_s \left[Ncdf(x, 0.4\mu, \sigma) - Ncdf(x, \mu, \sigma) \right] \qquad (7)$$

Again n_s is a coefficient to normalize the cumulative difference function. The difference in cumulative difference functions for the normal curve is found on some calculators (e.g., TI-83).

When this curve is also plotted in Figure 2, it is very close to the previous curves. Therefore, this presents a simple and apparently accurate method for estimating the exposure distribution.

CORRECTION FOR NIGHTTIME RADON ELEVATION

Radon testers with continuous monitors know that the late-night and early-morning readings should be somewhat higher than the late-afternoon and early-evening readings. It is one way to assess the validity of radon measurements. Hobbs[9] discussed this in detail. Most people are in their residences at night, just when indoor radon concentrations are the highest.

I use a femto-TECH RS410F radon survey device that remembers the number of radon counts each half hour. I looked at my last 39 measurements, corrected for background, and found the following: The average of the last 156 measurements between 5 and 6 AM was 3.42 pCi/L while the last 156 measurements between 5 and 6 PM was 2.08 pCi/L. Some houses were higher than others so, to account for this, I calculated the ratio of the measurements for each of the 39 houses. The average of all these ratios is 1.66, which is the factor I used to increase the California geometric mean for the exposure calculation in this paper.

By looking at the concentration ratio for specific houses, the amount of scatter in the data is reduced. The 95% confidence interval for the ratio is 1.45 to 1.89. In my area, there is no doubt that the stack effect is the primary driving force bringing in radon with soil gases during the nighttime hours.

CONCLUSIONS

In assessing the increased cancer risk from radon, the time-average exposure is a more meaningful measure than average concentration. Let me give a personal extreme example. There was tremendous concern in Santa Barbara, California, because one school building had several rooms with school-year averages above 4.0 pCi/L. One room had an average above 10 pCi/L. Every night the students went home and the room was shut up. During the night, the radon concentration in the room would sky rocket to magnitudes grater than 40 pCi/L. When the heating and ventilation came on in the morning the radon would dissipate. I never recorded a level during the school day that was greater than 1.0 pCi/L. We did dutifully install a mitigation system for the building that kept the average concentration in all rooms to negligible levels. But clearly, there was not much human exposure to begin with.

To assess the exposure in a residence, a rule of thumb is that there is a 75% occupancy rate. The data for California shows a 72% rate, and it is reasonable to expect that other parts of the USA would be higher. There is large scatter, however, in residence occupancy rate. As I showed above, a flat distribution from about 40% to 100% is a reasonable approximation for the California activity profiles. The time-average exposure 0.73 pCi/L is about 68% of the average residence concentration 1.08 pCi/L.

The average residence concentration itself has been increased because of the nighttime stack effect. This, as the schoolroom example clearly shows, violates the assumption of independence between average radon concentration and activity patterns. If you go to work during the daytime and sleep at home at night, your effective radon exposure may be greatly increased.

REFERENCES

1. K.S. Liu, S.B. Hayward, J.R. Girman, B.A. Moed, and F.Y. Huang, "Annual Average Radon Concentrations in California Residences," *J Air & Waste Management Assoc 41*, pp. 1207-1212, September 1991.

2. D. Quinton, *California Statewide Radon Survey Screening Results,* jointly sponsored by US Environmental Protection Agency and the California Dept of Health Services, Sacramento, California, 1993.

3. W.E. Hobbs and L.Y. Maeda, "Identification and Assessment of a Small Geologically Localized Radon Hot Spot," *Environment International 22,* pp. S809-S817, January 1996.

4. A.V. Nero, M.B. Schwehr, W.W. Nazaroff, and K.L. Revzan, "Distribution of Airborne Radon-222 Concentrations in U.S. Homes," *Science 234,* pp. 992-997, 1986.

5. P.L. Jenkins, T.J. Phillips, E.J. Mulberg, and S.P. Hui, "Activity Patterns of Californians: Use of and Proximity to Indoor Pollutant Sources," *Atmospheric Environment 26A,* pp. 2141-2148, December 1992.

6. T.J. Phillips, P.L. Jenkins, and E.J. Mulberg, "Children in California: Activity Patterns and Presence of Pollutant Sources," Proceedings of the 84th Annual Meeting of the Air and Waste Management Association, Pittsburgh, PA, 1991.

7. M.D. Koontz, W.C. Evans, and C.R. Wilkes, *Development of a Model for Assessing Indoor Exposure to Air Pollutants, Final Report A933-157,* California Air Resources Board Research Division, Sacramento, California, January 1998.

8. Theorems for working lognormal distributions are similar to those for the normal distributions except that multiplication is substituted for addition. In this paper the parameters are designated the natural log of the geometric mean and the natural log of the standard deviation. The differentials are always assumed for the probability functions: if $z = kx$, $f(z)dz = f(kx)d(kx) = k\ f(kx)dx$.

9. W.E. Hobbs, "Ventilation Level from Time Profile of Radon," *1999 International Radon Symposium,* paper 12, Sponsored by AARST, Las Vegas, NV, November 1999.

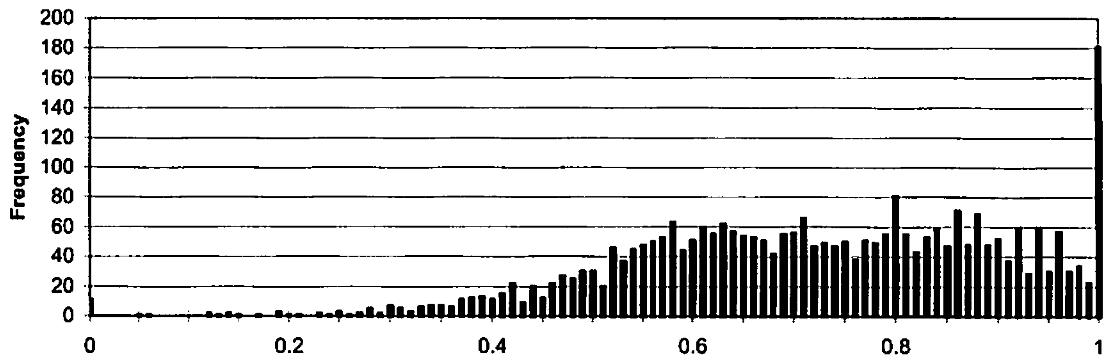

Figure 1. Percentiles for a sample of 2962 Californians

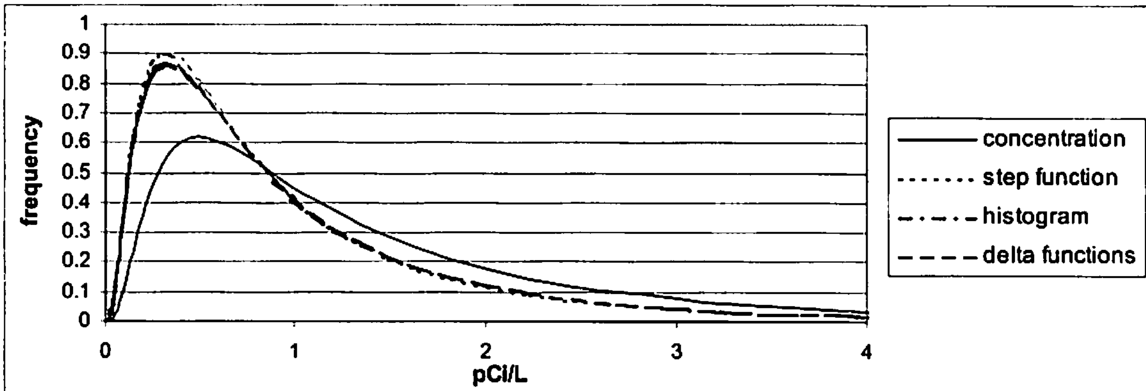

Figure 2. Comparison of time-average human exposure to radon using different activity approximations with the residential radon concentrations.

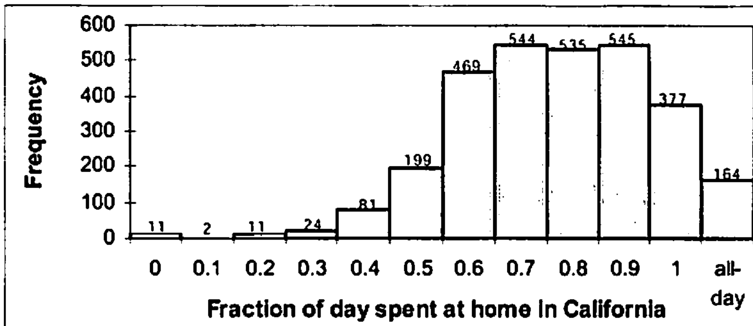

Figure 3. Histogram of activity profiles for sample of 2962 Californians

Appendix
Calculation of Residential Radon Exposure by Convolution of Home Concentration and Fraction of Day at Home

We proceed by first constructing the cumulative distribution of exposure $E(z)$ which is a expression giving the total probability that a persons exposure is less than or equal to z. The distributions $c(x)$ and $a(y)$ are assumed to be independent so that the product $c(x)\bullet a(y)$ is the probability that a person is California has a home with concentration x and is present in the building fraction y of the day. (A clear fallacy here is that some pollutants are the result of human activity in the building and, thus, larger values of y would correlate with larger values of x.)

A general formula for the cumulative exposure distribution is, by definition,

$$E(z) = \iint_{xy<z} c(x) \bullet a(y)\, dx\, dy \qquad (1)$$

This can be rewritten in two ways,

$$E(z) = \int_{x=0}^{\infty} \int_{y=0}^{z/x} c(x) \bullet a(y)\, dy\, dx = \int_{y=0}^{1} \int_{x=0}^{z/y} c(x) \bullet a(y)\, dx\, dy \qquad (2)$$

The probabilities are summed over the appropriate portion of the first quadrant. Note that because of the their assumed independence, we can write

$$E(z) = \int_{x=0}^{\infty} c(x) \int_{y=0}^{z/x} a(y)\, dy\, dx = \int_{y=0}^{1} a(y) \int_{x=0}^{z/y} c(x)\, dx\, dy \qquad (3)$$

The individual cumulative probability distributions for the concentrations and day fractions are evident in this formulation. But they are not evaluated since this usually doesn't help.

Let $y'=y\bullet x$ in the first expression and $x'=x\bullet y$ in the second expression,

$$E(z) = \int_{x=0}^{\infty} c(x) \int_{y'=0}^{z} a\left(\frac{y'}{x}\right) \frac{dy'}{x}\, dx = \int_{y=0}^{1} a(y) \int_{x'=0}^{z} c\left(\frac{x'}{y}\right) \frac{dx'}{y}\, dy \qquad (4)$$

The probability density function is found by differentiating with respect to z

$$e(z) = \frac{dE}{dz} = \int_{x=0}^{\infty} c(x) a\left(\frac{z}{x}\right) \frac{dx}{x} = \int_{y=0}^{1} a(y) c\left(\frac{z}{y}\right) \frac{dy}{y} \qquad (5)$$

The alternate forms appear quite different, but are equivalent. If the two distributions $a(y)$ and $c(x)$ are normalized, then $e(z)$ will also be normalized.

Suppose the activity at-home distribution $a(y)$ is a normalized step function between fractions b and d ($0 \le b < d \le 1$) designated $S(y,b,d)$. The population is assumed to have uniform probability of being home any fraction y between b and d. Since this function is zero outside of its limits, the above integrals can be simplified. This may be generalized since any function can be approximated to arbitrary accuracy by a histogram.

$$e(z) = \int_{x=z/b}^{z/d} \frac{c(x)}{x}(-dx) = \int_{x=z/d}^{z/b} \frac{c(x)}{x} dx = \int_{y=b}^{d} \frac{c(z/y)}{y} dy \qquad (6)$$

To check that the last two forms are identical, use $xy=z$. Since $c(x)$ is a well behaved probability function, the middle integral is easiest to evaluate.

The concentration distribution for many pollutants is extremely well approximated by a lognormal distribution

$$c(x)\, dx = \frac{\exp\left[-(\ln x - \mu)^2/(2\sigma^2)\right]}{x\sigma\sqrt{2\pi}} dx \qquad (7)$$

Where μ and σ are the natural logarithms of the geometric mean and standard deviation of the distribution.

Dividing by x

$$\begin{aligned}\frac{c(x)\,dx}{x} &= \frac{\exp\left[-(\ln x-\mu)^2/(2\sigma^2)\right]e^{-\ln x}}{\sigma\sqrt{2\pi}}\,d(\ln x)\\ &= \frac{\exp\left\{-\left[(\ln x)^2-2\mu\ln x+\mu^2+2\sigma^2\ln x\right]/(2\sigma^2)\right\}}{\sigma\sqrt{2\pi}}\,d(\ln x)\\ &= \frac{\exp\left\{-\left[(\ln x)^2-2(\mu-\sigma^2)\ln x+\mu^2\right]/(2\sigma^2)\right\}}{\sigma\sqrt{2\pi}}\,d(\ln x)\end{aligned} \qquad (8)$$

Completing the square,

$$\begin{aligned}\frac{c(x)\,dx}{x} &= e^{-\mu+\sigma^2/2}\,\frac{\exp\left\{-\left[(\ln x)^2-2(\mu-\sigma^2)\ln x+\mu^2-2\mu\sigma^2+\sigma^4\right]/(2\sigma^2)\right\}}{\sigma\sqrt{2\pi}}\,d(\ln x)\\ &= e^{-\mu+\sigma^2/2}\,\frac{\exp\left[-\left(\ln x-\mu+\sigma^2\right)^2/(2\sigma^2)\right]}{\sigma\sqrt{2\pi}}\,d(\ln x)\end{aligned} \qquad (9)$$

The expectation value for $1/x$ is simply the factor $\exp(-\mu+\sigma^2/2)$. Make the change of variable $\ln w = \ln x+\sigma^2$

$$\int_{z/d}^{z/b}\frac{c(x)\,dx}{x} = \frac{e^{-\mu+\sigma^2/2}}{\sigma\sqrt{2\pi}}\int_{w=zf/d}^{zf/b}\exp\left[-(\ln w-\mu)^2/(2\sigma^2)\right]\frac{dw}{w} \qquad (10)$$

Where the factor f=$\exp(\sigma^2)$. Note that this is not the square of the geometric standard deviation, but rather the exponential of the square of the logarithm of the geometric standard deviation. For example, if the geometric standard deviation was e^{σ}=1.5, then σ=0.4 and $\exp(\sigma^2)\approx 1.17$, not $(e^{\sigma})^2$=2.25. The important result is that this final integral is nothing more than an error function, or the cumulative probability of a normal distribution.

$$\int_{z/d}^{z/b}\frac{c(x)\,dx}{x} = e^{-\mu+\sigma^2/2}\left\{Ncdf\left[\mu,\sigma,\ln(zf/b)\right]-Ncdf\left[\mu,\sigma,\ln(zf/d)\right]\right\} \qquad (11)$$

Note that after completing the square (Eqn. 9), the integrand is simply a lognormal function with the mean shifted σ^2 units to the left. We have effectively just proved that

$$\text{lognormal}(x,k\mu,\sigma) = \text{lognormal}(x/k.\mu,\sigma) \qquad (12)$$

ACKNOWLEDGEMENT: J. Edmondson of SBCC turned me on to ***Modern Probability Theory and Its Applications*** by Parzen. The mathematics of this appendix is discussed more fully there.