

ENERGY STAR® Performance Ratings Technical Methodology for K-12 School

This document presents specific details on the EPA's analytical result and rating methodology for K-12 School. For background on the technical approach to development of the Energy Performance Ratings, refer to *Energy Performance Ratings – Technical Methodology* (http://www.energystar.gov/ia/business/evaluate_performance/General_Overview_tech_methodology.pdf).

Model Release Date¹

Most Recent Update: February 2009

Previous Update: January 2004

Original Release Date: April 2000

Portfolio Manager Definitions

K-12 School applies to facility space used as a school building for Kindergarten through 12th grade students. This does not include college or university classroom facilities and laboratories, vocational, technical, or trade schools. The total gross floor area should include all supporting functions such as administrative space, conference rooms, kitchens used by staff, lobbies, cafeterias, gymnasiums, auditoria, laboratory classrooms, portable classrooms, greenhouses, stairways, atria, elevator shafts, small landscaping sheds, storage areas, etc.

Reference Data

The K-12 School regression model is based on data from the Department of Energy, Energy Information Administration's 2003 Commercial Building Energy Consumption Survey (CBECS). Detailed information on this survey, including complete data files, is publicly available at: <http://www.eia.doe.gov/emeu/cbecs/contents.html>.

Data Filters

Four types of filters are applied to define the peer group for comparison and to overcome any technical limitations in the data: Building Type Filters, EPA Program Filters, Data Limitation Filters, and Analytical Filters. A complete description of each of these categories is provided in Section V of the general technical description document: *Energy Performance Ratings – Technical Methodology*. **Table 1** presents a summary of each filter applied in the development of the K-12 School model, the rationale behind the filter, and the resulting number of observations after the filter is applied. After all filters are applied, the remaining data set has 353 observations.

¹ Periodic updates to the model occur to reflect the most current available market data. The original model was developed using the CBECS 1995 database; a previous update in 2004 reflected the CBECS 1999 database. The most current update of February 2009 reflects the CBECS 2003 database.

The reasons for applying filters on the use and quantity of propane are worthy of additional discussion. In CBECS, major fuel use is reported in exact quantities of consumption. However, if a building uses propane, the amount of propane is reported according to the variable PRAMT8, which uses ranges rather than exact quantities (e.g. less than 100 gallons, 100 to 500 gallons, etc). Therefore, the quantity must be estimated within the range. To limit error associated with this estimation, EPA applies two restrictions to the propane quantity.

1. The quantity of propane expressed by PRAMT8 must be 1000 gallons or smaller.
2. The value of propane cannot account for more than 10% of the total source energy use. Because the exact quantity of propane is not reported, this cap ensures that the quantity of propane entered will not introduce undue error into the calculation of total energy consumption. In order to apply this 10% limitation, the value at the high end of the propane category is employed (e.g. for the category of less than 100, a value of 99 is used). If the 10% cap is not exceeded, then EPA will use the value at the middle of the range to calculate total energy use (e.g. for the category of less than 100, a value of 50 is used).

Table 1 Summary of K-12 School Model Filters		
Condition for Including an Observation in the Analysis	Rationale	Number Remaining
PBAPLUS8= 28 or 29	Building Filter – CBECS defines building types according to the variable “PBAPLUS8.” Elementary/middle schools are coded as PBAPLUS8 = 28 and High Schools are coded as PBAPLUS8 = 29.	456
Must operate for at least 30 hours per week	EPA Program Filter – Baseline condition for being a full time K-12 School.	448
Must have energy consumption data	EPA Program Filter – Baseline condition for being a full time K-12 School.	447
Must operate for at least 8 months per year	EPA Program Filter – Baseline condition for being a full time K-12 School.	434
A single activity must characterize greater than 50% of the floor space ²	EPA Program Filter – In order to be considered part of the K-12 School peer group, more than 50% of the building must be defined by one of those activities.	432
Must have square foot less than or equal to 1,000,000	Data Limitation Filter – CBECS masks actual values above 1,000,000 using regional averages.	431
If propane is used, the amount category (PRAMTC8) must equal 1, 2, or 3	Data Limitation Filter – Cannot estimate propane use if the quantity is “greater than 1000” or unknown.	408
If propane is used, the maximum estimated propane amount must be 10% or less of the total source energy	Data Limitation Filter – Because propane values are estimated from a range, propane is restricted to 10% of the total source energy.	408
Must not use chilled water	Data Limitation Filter – CBECS does not collect quantities of chilled water	400
Must have square foot greater than or equal to 5,000	Analytical Filter – Analysis could not model behavior for buildings smaller than 5,000 ft ² .	361
Must have Source EUI less than 450 kbtu/ft ²	Analytical Filter – Values determined to be statistical outliers.	353

Dependent Variable

The dependent variable in the K-12 School analysis is source energy use intensity (source EUI). Source EUI is equal to the total source energy use of the facility divided by the gross floor area. By setting source EUI as the dependent variable, the regressions analyze the key drivers of source EUI – those factors that explain the variation in source energy per square foot in K-12 Schools.

² This filter is applied by a set of screens. If the variable ONEACT8=1, this indicates that one activity occupies 75% or more of the building. If the variable ONEACT8=2, then the building can specify up to 3 activities (ACT18, ACT28, ACT38). One of these activities must be Education (PBAX8=16) and must account for more than 50% of the floor area.

Independent Variables

General Overview:

The CBECS data contains numerous building operation questions that EPA identified as potentially important for K-12 Schools. Based on a review of the available variables in the CBECS data, in accordance with the EPA criteria for inclusion³, EPA analyzed the following variables⁴:

- SQFT8 – Square footage
- WKHRS8 – Weekly hours of operation
- EDSEAT8 – Student seating capacity
- PCNUM8 – Number of personal computers
- SRVNUM8 – Number of servers
- PRNTRN8 – Number of printers
- COPRN8 – Number of photocopiers
- RFGWIN8 – Number of walk-in refrigeration units
- RFGOPN8 – Number of open refrigerated cases
- RFGRSN8 – Number of residential refrigerators
- RFGCLN8 – Number of closed refrigerated cases
- RFGVNN8 – Number of refrigerated vending machines
- COOK8 – Energy used for cooking (yes/no)
- FDRM8 – Commercial food preparation area (yes/no)
- CAF8 – Cafeteria or large restaurant (yes/no)
- ELEVTR8 – Elevators (yes/no)
- NELVTR8 – Number of elevators
- NFLOOR8 – Number of floors
- MONUSE8 – Months of year in use
- OPNWE8 – Open weekends (yes/no)
- POOL8 – Indoor pool (yes/no)
- HTPOOL8 – Heated pool (yes/no)
- YRCON – Year of construction
- RENOV8 – Renovations since 1980 (yes/no)
- HEATP8 – Percent heated
- COOLP8 – Percent cooled
- HDD658 – Heating degree days
- CDD658 – Cooling degree days

EPA performed extensive review on all of these operational characteristics. In addition to reviewing each characteristic individually, characteristics were reviewed in combination with each other (e.g., Heating Degree Days * Percent Heated). As part of the analysis, some variables were reformatted to reflect the physical relationships of building components. For example, the

³ For a complete explanation of these criteria, refer to *Energy Performance Ratings – Technical Methodology* (http://www.energystar.gov/ia/business/evaluate_performance/General_Overview_tech_methodology.pdf).

⁴ Note that the 8 at the end of all variables indicates that the 2003 CBECS survey is the eighth survey conducted by the Energy Information Administration

number of personal computers is typically evaluated in a density format. The number of computers *per square foot* (not the gross number of computers) is expected to be correlated with the energy use per square foot. In addition, based on analytical results and residual plots, variables were examined using different transformations (such as the natural logarithm). The analysis consisted of multiple regression formulations. These analyses were structured to find the combination of statistically significant operating characteristics that explained the greatest amount of variance in the dependent variable: source EUI.

Based on the K-12 School regression analysis, the following seven characteristics were identified as key explanatory variables that can be used to estimate the expected average source EUI (kBtu/ft²) in K-12 Schools:

- Natural log of gross square feet
- Whether the school is open on weekends (1 = yes, 0 = no)
- Number of walk-in refrigerators per 1,000 square feet
- Whether there is energy used for cooking (1 = yes, 0 = no)
- Number of personal computers (PCs) per 1,000 square feet
- (Natural log of heating degree days) times Percent of the building that is heated
- (Natural log of cooling degree days) times Percent of the building that is cooled

High School analysis:

In addition, analysis revealed that high school buildings have different responses to gross square feet and cooling degrees than elementary/middle schools. Due to this unique response, the final regression includes a dummy variable to account for the different average energy use at High Schools, in addition to two interactive terms to account for the different impact that size and cooling degree days have on high schools. These variables are as follows:

- Yes/No variable indicating whether the building is a high school
- Additional floor area adjustment
- Additional climate adjustments
 - Cooling degree days times Percent of the building that is cooled
 - (Natural log of cooling degree days) times percent of the building that is cooled

The determination of these adjustments was based on a substantial analysis of the data and the differences among types of K-12 schools. EPA investigated a wide variety of regression formulations. The adjustments for high schools described above were determined to be statistically significant when added to the K-12 Schools regression model. These adjustments improved the overall significance of the K-12 Schools regression mode, and resulted in more equitable energy performance ratings for both elementary/middle schools and high schools.

Model Testing:

Finally, once the final regression model was developed EPA performed a variety of test runs using existing K-12 School buildings that have been entered in Portfolio Manager. This existing data provided another set of buildings to examine in addition to the CBECS data, to determine

the average ratings and distributions, and to assess the impacts and adjustments. This analysis provided a second level of confirmation that the final regression model produces robust results that are unbiased with respect to the key operational characteristics such as building size, computer density, and heating and cooling degree days.

It is important to reiterate that the final regression model is based on the nationally representative CBECS data, not data previously entered into EPA's Portfolio Manager.

Regression Modeling Results

The final regression is a weighted ordinary least squares regression across the filtered data set of 353 observations. The dependent variable is source EUI. Each independent variable is centered relative to the mean value, presented in **Table 2**. The final model is presented in **Table 3**. All model variables are significant at the 95% confidence level or better, as shown by the significance levels (a p-level of less than 0.05 indicates 95% confidence), with the exception of the high school yes/no variable, which has a lower level of significance (42%). The high school variable is unique given the significant interaction terms for high school with cooling degree days and floor area. Because these interaction terms are highly significant, it is standard practice in statistical analysis to retain the base dummy variable (high school yes/no) in the regression model. This practice provides a more appropriate intercept for the high school population, distinguishing it from the other schools.

The model has an R^2 value of 0.268, indicating that this model explains 26.8% of the variance in source EUI for K-12 School buildings. Because the final model is structured with energy per square foot as the dependent variable, the explanatory power of square foot is not included in the R^2 value, thus this value appears artificially low. Re-computing the R^2 value in units of source energy⁵, demonstrates that the model actually explains 83.8% of the variation of source energy of K-12 Schools. This is an excellent result for a statistically based energy model.

Detailed information on the ordinary least squares regression approach, the methodology for performing weather adjustments, and the independent variable centering technique is available in the technical document: *Energy Performance Ratings – Technical Methodology*.

⁵ The R^2 value in Source Energy is calculated as: $1 - (\text{Residual Variation of Y}) / (\text{Total Variation of Y})$. The residual variation is sum of $(\text{Actual Source Energy}_i - \text{Predicted Source Energy}_i)^2$ across all observations. The Total variation of Y is the sum of $(\text{Actual Source Energy}_i - \text{Mean Source Energy})^2$ across all observations.

Table 2
Descriptive Statistics for Variables in Final Regression Model

Variable	Full Name	Mean	Minimum	Maximum
SrcEUI	Source Energy per Square Foot	151.6	11.22	400.0
Highsch	High School (yes/no)	0.2298	0.000	1.000
LNSqFt	Natural Log of Square Foot	10.20	8.517	13.10
SqFt	Square Feet	47310	5000	490000
Opnwe	Open Weekends	0.2730	0.000	1.000
WalkinDen	Number of Walk-in Refrigerators per 1000 ft ²	0.0109	0.000	0.1928
Cook	Presence of Cooking	0.5358	0.000	1.000
PCDen	Number of Computers per 1000 ft ²	1.742	0.000	9.537
CDDxPC	Cooling Degree Days x Percent Cooled	1316	0.000	5064
LNHDDxPH	(Natural Log of Heating Degree Days) x Percent Heated	7.716	0.2821	9.139
LNCDDxPC	(Natural Log of Cooling Degree Days) x Percent Cooled	5.045	0.000	8.530

Note:

- Statistics are computed over the filtered data set (n=353 observations).
- Values are weighted by the CBECS variable ADJWT8.
- The mean values are used to center variables for the regression.

Table 3				
Final Regression Modeling Results				
Dependent Variable	Source Energy Intensity (kBtu/ft ²)			
Number of Observations in Analysis	353			
Model R ² value	0.268			
Model F Statistic	11.35			
Model Significance (p-level)	0.000			
	Unstandardized Coefficients	Standard Error	T value	Significance (p-level)
(Constant)	131.9	6.668	19.78	0.0000
Higsch	4.377	7.922	0.5526	0.5809
C_LNHDDxPH	8.974	2.582	3.476	0.0005
C_LNCDDxPC	6.389	1.230	5.193	0.0000
C_LNSqFt	-19.26	4.295	-4.484	0.0000
Opnwe	18.43	7.488	2.461	0.0143
C_WalkinDen	574.7	150.1	3.830	0.0002
Cook	24.20	9.416	2.570	0.0106
C_PCDen	9.568	2.336	4.096	0.0001
higschxC_SqFt	0.00021	0.0001	2.490	0.0133
higschxC_CDDxPC	0.0285	0.0093	3.071	0.0023
higschxC_LNCDDxPC	-11.75	3.781	-3.107	0.0020
<i>Note:</i>				
- The regression is a weighted ordinary least squares regression, weighted by the CBECS variable "ADJW8".				
- The prefix C_ on each variable indicates that it is centered. The centered variable is equal to difference between the actual value and the observed mean. The observed mean values are presented in Table 2 .				
- Full variable names and definitions are presented in Table 2 .				
- Unlike other variables, the yes/no variables (Cook and OpnWe) are not centered. The coefficient adjustments represent the adjustment for Schools that have these characteristics.				
- The Higsch terms are not centered because they represent a multiplier on the already centered variables C_Sqft, C_CDDxPC, and C_LNCDDxPC, hence the variable is computed as high school times square foot for high school.				

K-12 School Lookup Table

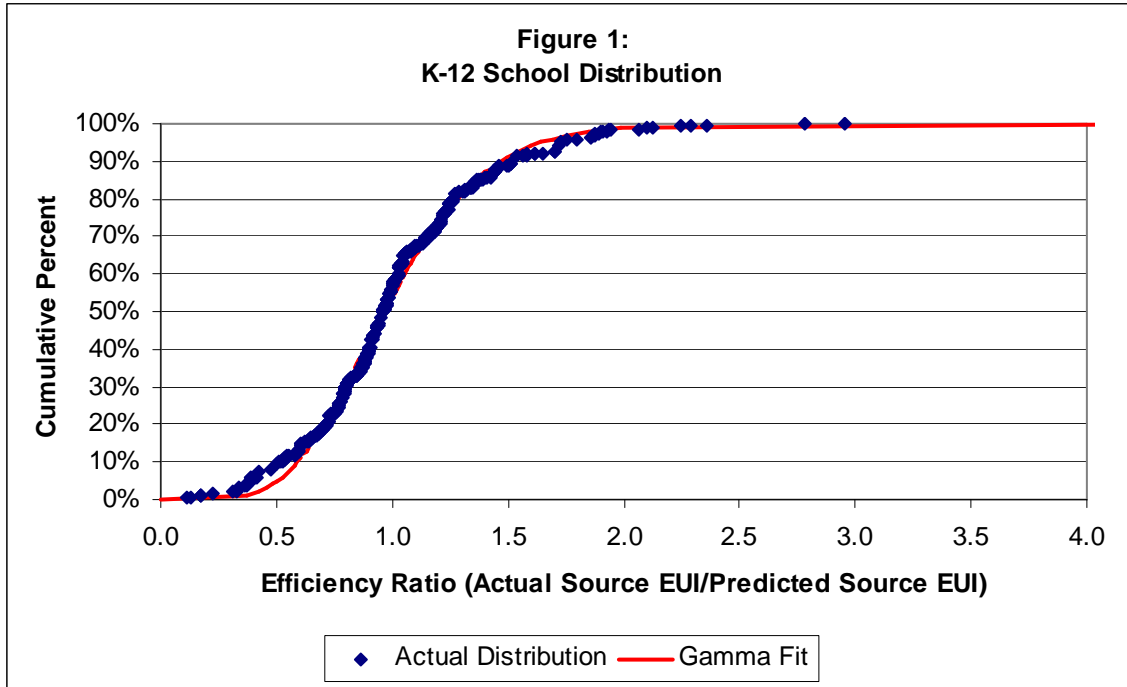
The final regression model (presented in **Table 3**) yields a prediction of source EUI based on a building's operating constraints. Some buildings in the CBECS data sample use more energy than predicted by the regression equation, while others use less. The *actual* source EUI of each CBECS observation is divided by its *predicted* source EUI to calculate an energy efficiency ratio:

$$\text{Energy Efficiency Ratio} = \text{Actual Source EUI} / \text{Predicted Source EUI}$$

A lower efficiency ratio indicates that a building uses less energy than predicted, and consequently is more efficient. A higher efficiency ratio indicates the opposite.

The efficiency ratios are sorted from smallest to largest and the cumulative percent of the population at each ratio is computed using the individual observation weights from the CBECS dataset. **Figure 1** presents a plot of this cumulative distribution. A smooth curve (shown in red) is fitted to the data using a two parameter gamma distribution. The fit is performed in order to

minimize the sum of squared differences between each building’s actual percent rank in the population and each building’s percent rank with the gamma solution. The final fit for the gamma curve yielded a shape parameter (alpha) of 8.22911 and a scale parameter (beta) of 0.12198. For this fit, the sum of the squared error is 0.17517.



The final gamma shape and scale parameters are then used to calculate the efficiency ratio at each percentile (1 to 100) along the curve. For example, the ratio on the gamma curve at 1% corresponds to a rating of 99; only 1% of the population has a ratio this small or smaller. The ratio on the gamma curve at the value of 25% is 0.7601. The ratio on the gamma curve at the value of 25% will correspond to the ratio for a rating of 75; only 25% of the population has ratios this small or smaller. The complete lookup table is presented at the end of the document. In order to read this lookup table, note that if the ratio is less than 0.3701 the rating for that building should be 100. If the ratio is greater than or equal to 0.3701 and less than 0.4212 the rating for the building should be 99, etc.

Example Calculation

As detailed in the document *Energy Performance Ratings – Technical Methodology*, there are five steps to compute a rating. The following is a specific example with the K-12 School model:

Step 1 – User enters building data into Portfolio Manager

For the purposes of this example, sample data is provided.

- Energy data
 - Total annual electricity = 700,000 kWh

- Total annual natural gas = 20,000 therms
- Note that this data is actually entered in monthly meter entries
- Operational data
 - Gross floor area (ft²) = 100,000
 - High School = Yes
 - Open Weekends = Yes
 - Walk-In Refrigerators = 0
 - Cook = No
 - Number of personal computers = 200
 - Percent heated = 100
 - Percent cooled = 100
 - HDD (provided by Portfolio Manager, based on zip code) = 4937
 - CDD (provided by Portfolio Manager, based on zip code) = 1046

Step 2 – Portfolio Manager computes the Actual Source Energy Use Intensity

In order to compute actual source EUI, Portfolio Manager must convert each fuel from the specified units (e.g. kWh) into Site kBtu, and must convert from Site kBtu to Source kBtu.

- Convert the meter data entries into site kBtu
 - Electricity: (700,000 kWh)*(3.412kBtu/kWh) = 2,388,400 kBtu Site
 - Natural gas: (20,000 therms)*(100kBtu/therm) = 2,000,000 kBtu Site
- Apply the source-site ratios to compute the source energy
 - Electricity:
2,388,400 Site kBtu*(3.34 Source kBtu/Site kBtu) = 7,977,256 kBtu Source
 - Natural Gas:
2,000,000 Site kBtu *(1.047 Source kBtu/Site kBtu) = 2,094,000 kBtu Source
- Combine source kBtu across all fuels
 - 7,977,256 kBtu + 2,094,000 kBtu = 10,071,256 kBtu
- Divide total source energy by gross floor area
 - Source EUI = 10,071,256 kBtu/100,000ft² = 100.7 kBtu/ft²

Step 3 – Portfolio Manager computes the Predicted Source Energy Intensity

Portfolio Manager uses the building data entered under Step 1 to compute centered values for each operating parameter. These centered values are entered into the K-12 School regression equation to obtain a predicted source EUI.

- Calculate centered variables
 - Use the operating characteristic values to compute each variable in the model.
(e.g. LN(Square Foot) = LN(100,000) = 11.51)
 - Subtract the reference centering value from calculated variable
(e.g. LN(Square Foot) – 10.20 = 11.51 – 10.20 = 1.31)
 - These calculations are summarized in **Table 4**
- Compute predicted source energy use intensity
 - Multiply each centered variable by the corresponding coefficient in the model
(e.g. Coefficient*CenteredLN(Square Foot) = -19.26*1.31 = -25.23)
 - Take the sum of these products (i.e. coefficient*CenteredVariable) and add to the constant (this yields a predicted Source EUI of 125.9 kBtu/ft²)

- This calculation is summarized in **Table 5**

Step 4 – Portfolio Manager computes the energy efficiency ratio

The energy efficiency ratio is equal to: Actual Source EUI/ Predicted Source EUI

- Ratio = $100.7/125.9 = 0.7998$

Step 5 – Portfolio Manager looks up the efficiency ratio in the lookup table

Starting at 100 and working down, Portfolio Manager searches the lookup table for the first ratio value that is larger than the computed ratio for the building.

- A ratio of 0.7998 is less than 0.8039 (requirement for 70) but greater than 0.7953 (requirement for 71)
- ***The rating is 70***

Table 4 Example Calculation – Computing Building Centered Variables				
Operating Characteristic	Formula to Compute Variable	Building Variable Value	Reference Centering Value	Building Centered Variable (Variable Value - Center Value)
Highsch	Highsch	1.000	NA	1.000
C_LNHDDxPH	LN(HDD)*Percent Heated	8.505	7.716	0.7890
C_LNCDDxPC	LN(CDD)*Percent Cooled	6.953	5.045	1.908
C_LNSqFt	LN(Square Foot)	11.51	10.20	1.310
Opnwe	Opnwe	1.000	NA	1.000
C_WalkinDen	# Walk-ins/ft ² *1000	0.000	0.0109	-0.0109
Cook	Cook	0.0000	NA	0.000
C_PCDen	#Computers/ft ² *1000	2.000	1.742	0.2580
C_SqFt	Square Feet	100000	47310	52690
C_CDDxPC	CDD*Percent Cooled	1046	1316	-270.0
highschxC_SqFt	Highsch * C_SQFT	52690	NA	52690
highschxC_CDDxPC	Highsch * C_CDDxPC	-270.0	NA	-270.0
highschxC_LNCDDxPC	Highsch * C_LN(CDD)xPC	1.908	NA	1.908
<i>Note</i> <ul style="list-style-type: none"> - Densities are always expressed as the number per 1,000 square feet. - The center reference values are the weighted mean values from the CBECS population, show in Table 2. - The Highsch terms are not centered because they represent a multiplier on the already centered variables C_Sqft, C_CDDxPC, and C_LNCDDxPC, hence the variable is computed as high school times square foot for high school. In this example the building is a high school, so the value is 1 times each of the centered terms. If this were not a high school, the values would all be zero. 				

Table 5 Example Calculation – Computing predicted Source EUI			
Operating Characteristic	Centered Variable	Coefficient	Coefficient * Centered Variable
Constant	NA	131.9	131.9
Highsch	1	4.377	4.377
C_LNHDDxPH	0.7890	8.974	7.080
C_LNCDDxPC	1.908	6.389	12.19
C_LNSqFt	1.310	-19.26	-25.23
Opnwe	1	18.43	18.43
C_WalkinDen	-0.0109	574.7	-6.264
Cook	0	24.20	0
C_PCDen	0.2580	9.568	2.469
highschxC_SqFt	52690	0.00021	11.06
highschxC_CDDxPC	-270.0	0.0285	-7.695
highschxC_LNCDDxPC	1.908	-11.75	-22.42
Predicted Source EUI (kBtu/ft²)			125.9

Attachment

Table 6 lists the energy efficiency ratio cut-off point for each rating, from 1 to 100.

Table 6 Lookup Table for K-12 School Rating					
Rating	Cumulative Percent	Ratio	Rating	Cumulative Percent	Ratio
100	0%	<0.3710	50	50%	0.9720
99	1%	0.4212	49	51%	0.9807
98	2%	0.4555	48	52%	0.9894
97	3%	0.4825	47	53%	0.9982
96	4%	0.5054	46	54%	1.0070
95	5%	0.5254	45	55%	1.0159
94	6%	0.5434	44	56%	1.0250
93	7%	0.5599	43	57%	1.0341
92	8%	0.5752	42	58%	1.0433
91	9%	0.5895	41	59%	1.0526
90	10%	0.6030	40	60%	1.0620
89	11%	0.6159	39	61%	1.0716
88	12%	0.6283	38	62%	1.0813
87	13%	0.6401	37	63%	1.0911
86	14%	0.6515	36	64%	1.1011
85	15%	0.6626	35	65%	1.1113
84	16%	0.6733	34	66%	1.1217
83	17%	0.6837	33	67%	1.1322
82	18%	0.6939	32	68%	1.1429
81	19%	0.7039	31	69%	1.1539
80	20%	0.7137	30	70%	1.1651
79	21%	0.7232	29	71%	1.1765
78	22%	0.7327	28	72%	1.1883
77	23%	0.7419	27	73%	1.2003
76	24%	0.7511	26	74%	1.2126
75	25%	0.7601	25	75%	1.2253
74	26%	0.7690	24	76%	1.2384
73	27%	0.7779	23	77%	1.2519
72	28%	0.7866	22	78%	1.2659
71	29%	0.7953	21	79%	1.2803
70	30%	0.8039	20	80%	1.2953
69	31%	0.8124	19	81%	1.3110
68	32%	0.8209	18	82%	1.3273
67	33%	0.8293	17	83%	1.3444
66	34%	0.8378	16	84%	1.3623
65	35%	0.8461	15	85%	1.3812
64	36%	0.8545	14	86%	1.4013
63	37%	0.8628	13	87%	1.4226
62	38%	0.8712	12	88%	1.4455
61	39%	0.8795	11	89%	1.4702
60	40%	0.8878	10	90%	1.4970
59	41%	0.8961	9	91%	1.5266
58	42%	0.9044	8	92%	1.5595
57	43%	0.9128	7	93%	1.5967
56	44%	0.9211	6	94%	1.6400
55	45%	0.9295	5	95%	1.6917
54	46%	0.9380	4	96%	1.7567
53	47%	0.9464	3	97%	1.8455
52	48%	0.9549	2	98%	1.9911
51	49%	0.9634	1	99%	>=1.9911