

Appendix

Random Variables and Probability Distributions

Random Variables

This section includes discussion on random variables types and characteristics of discrete and continuous distributions. It also introduces the uniform (rectangular) distribution. Technology will not help you with this material.

Means and Variances of Random Variables

Calculating the mean and variance of a discrete random variable can be done using software as a calculator.



1. Enter the values of the variable in one column and the probabilities in a second.
2. Place the cursor in a blank column adjacent to one of your columns.
3. Enter a formula to multiply the probability and the value, for example, $=A1*B1$.
4. Drag the cursor to copy the formula down the rows.
5. Place the cursor in an empty cell and enter a formula to sum the entries just created to find the mean, for example, $=sum(C1..Cn)$
6. Calculate the variance $\sigma^2 = \sum(x_i - \mu)^2 p_i$ by once again placing the cursor in an empty cell (for example, D1) and entering the formula (assuming the mean is in cell F1) $=(A1-F$1)^2*B1$
7. Sum the entries just created as was done in step 5.



1. Enter the values of the variable in one column and the probabilities in a second.
2. Click **Analyze** → **Distribution**.
3. Click to enter the column with the variable values into Y and the column with the probabilities into Weight.
4. Click **OK**.

In the results, the mean given is μ , and the standard deviation is σ . Square σ to find the variance.



Enter the values of the variable in one column (say, C1) and the probabilities in a second (say, C2.)

Computations are easiest done with commands, so click in the “Session” window and then click **View → Command line history**. In the command line box at right, enter the following sequence of commands. Click Run and use the cursor to highlight and delete the old command after each command.

```
Let C3=C1*C2
Sum C3 (This displays the mean.)
Let C4=(C1-mean)**2*C2 (where “mean” is the value just calculated)
Sum C4 (This displays the variance.)
```



1. Enter the values of the variable in one column and the probabilities in a second, say, X and P.

2. Click **Transform → Compute Variable**.

3. Enter a new variable name (say, XP) and the formula $X \cdot P$. Click **OK**.

4. Click **Analyze → Descriptive Statistics → Descriptives**.

5. Click to enter the variable just created (XP in this example) as the variable and then **Options**. Check the box next to “Sum.” Click **Continue → OK**. The entry for the sum is the mean of the random variable.

6. Click **Transform → Compute Variable**.

7. Enter a new variable name (say, XMUSQ) and the formula

$$(X - \text{mean}) ** 2 * P$$

where mean is the value just found. Click **OK**.

8. Click **Analyze → Descriptive Statistics → Descriptives**.

9. Click to enter the variable just created (XMUSQ in this example). The sum displayed is the variance.



CrunchIt! cannot perform these calculations; at this time, you cannot obtain the sum of a column (variable).



1. Enter the variable values in one list (say, L1) and the probabilities in another (say, L2).

2. Press **[STAT]** **[↓]** **[1]** (for 1-VarStats).

3. Assuming you are using the latest version of the operating system, enter L1 as the “List” and L2 as the “Frq.”

4. In the displayed results, \bar{x} is μ (the calculator only has one symbol for a mean), and σ_X is the population standard deviation.

5. For the variance (without rounding error), press **[VARS]** **[5]** (Statistics) **[4]** (for σ_X) **[x²]**.



Calculate the mean as shown in the example commands below.

```
> x<-c(1,2,3,4,5)
> p<-c(.1,.2,.2,.1,.4)
> xp=x*p
> sum(xp)
[1] 3.5
```

Calculate the variance as shown in the example commands below.

```
> xmmusq=(x-3.5)**2*p
> sum(xmmusq)
[1] 2.05
```

Distribution Calculation



Binomial Distribution

1. Click an empty cell in the spreadsheet.
2. Click **Formulas** → **More Functions** → **Statistical**.
3. Scroll down the list of functions and double-click on the “BINOM. DIST” function.
4. In the “Number_s” box, input the specified value of k (the number of successes).
5. In the “Trials” box, input the value of n (number of trials/observations).
6. In the “Probability s” box, input the value of p (probability of success).
7. If you wish to have $P(X = k)$, then input “0” or “FALSE” in the “Cumulative” box. If you wish to have $P(X \leq k)$, then input “1” or “TRUE” in the “Cumulative” box.
8. Click **OK** to find the results.

Poisson Distribution

1. Click an empty cell in the spreadsheet.
2. Select the Statistical function menu as described previously.
3. Scroll down the list of functions and double-click on the “POISSON. DIST” function choice.
4. In the “X” box, input the specified value of k (the number of successes).
5. In the “Mean” box, input the value of μ (mean number of successes per unit of measure).
6. If you wish to have $P(X = k)$, then input “0” “FALSE” in the “Cumulative” box. If you wish to have $P(X \leq k)$, then input “1” or “TRUE” in the “Cumulative” box.
7. Click **OK** to find the results.



Binomial Distribution

1. Click **Applets** → **Distribution Calculator**.
2. Select “Binomial” for the distribution.
3. Enter p (success) and n.
4. Use the radio buttons to select the type of calculation desired (input values to obtain a probability or enter a probability to obtain a value).
5. Select the probability option and enter the value(s) to look up in the bottom right section and press the enter key to see shading on distribution.

Poisson Distribution

1. Click **Applets** → **Distribution Calculator**.
2. Select “Poisson” for the distribution.
3. Enter the mean.
4. Select the type of calculation desired (input values to obtain a probability or enter a probability to obtain a value).
5. Select the probability option and enter the value(s) to look up in the bottom right section and press the enter key to see shading on distribution.

Note: In both of these calculations, to find, for example, $P(X = 5)$, use the probability option “ $q1 < X \leq q2$.” Enter 4 for $q1$ and 5 for $q2$.



Binomial Distribution

1. Click **Calc** → **Probability Distributions** → **Binomial**.
2. Use the radio buttons to choose “Probability” if you wish to find $P(X = k)$ or “Cumulative probability” if you wish to find $P(X \leq k)$.
3. Input the Number of trials (n) and the probability of success (p) in the “Event probability” box.
4. In most cases, you have a specific value k in mind; select “Input constant.” In the box next to this option, enter the specified value of k .
5. Click **OK** to find the results reported in the Session window.

Poisson Distribution

1. Click **Calc** → **Probability Distributions** → **Poisson**.
2. Use the radio buttons to choose “Probability” if you wish to find $P(X = k)$ or “Cumulative probability” if you wish to find $P(X \leq k)$.
3. Input the value of μ (mean number of successes per unit of measure) in the “Mean” box.
4. In most cases, you have a specific value k in mind; select “Input constant.” In the box next to this option, enter the specified value of k .
5. Click **OK** to find the results reported in the Session window.

In either case, to find a value given a cumulative probability (that is, to find k given $P(X \leq k)$), select the “Inverse cumulative probability” option.



Binomial Distribution

1. Click **Transform** → **Compute Variable**.
2. In “Target Variable,” enter a name for the new variable, then click in the “Numerical Expression” box.
3. In “Function Group” at right, select **PDF and Noncentral** for $P(X = k)$ and then **PDF.Binom**.
4. In “Function Group” at right, select **CDF and Noncentral** for $P(X \leq k)$ and then **CDF.Binom**.
5. In either case, enter the parameters: k , n (number of trials), and p (success).
6. A new column will be added to the data set containing the desired probability.

Poisson Distribution

1. Click **Transform** → **Compute Variable**.
2. In “Target Variable,” enter a name for the new variable, then click in the “Numerical Expression” box.
3. In “Function Group” at right, select **PDF and Noncentral** for $P(X = k)$ and then **PDF.Poisson**.
4. In “Function Group” at right, select **CDF and Noncentral** for $P(X \leq k)$ and then **CDF.Poisson**.
5. In either case, enter the parameters: k and μ .
6. A new column will be added to the data set containing the desired probability. SPSS does not have an inverse function for these distributions.

CRUNCH IT! Binominal Distribution

1. Click **Distribution Calculator** → **Binomial**.
2. Enter n and p .
3. Use the probability tab, then select the appropriate symbol ($=$, \leq , and so on).
4. Enter k .
5. Click **Calculate**.

Poisson Distribution

1. Click **Distribution Calculator** → **Poisson**.
2. Enter lambda (m).
3. Use the probability tab, then select the appropriate symbol ($=$, \leq , and so on).
4. Enter k .
5. Click “Calculate.”

In either case, to find a value given a cumulative probability (that is, to find k given $P(X \leq k)$), select the Quantile tab and enter $P(X \leq k)$, then click “Calculate.”



TI-83/-84

Binominal Distribution

1. For $P(X = k)$, press 2^{nd} [VAR]^{S} (DISTR) then arrow to `binompdf(` and press [ENTER] .
2. For $P(X \leq k)$, press 2^{nd} [VAR]^{S} (DISTR) then arrow to `binomcdf(` and press [ENTER] .
3. In both cases, enter n (“trials”), p , and the value to look up (“ x value”).
4. Press [ENTER] .

Poisson Distribution

1. For $P(X = k)$, press 2^{nd} [VAR]^{S} (DISTR) then arrow to `poissonpdf(` and press [ENTER] .
2. For $P(X \leq k)$, press 2^{nd} [VAR]^{S} (DISTR) then arrow to `poissoncdf(` and press [ENTER] .
3. In both cases, mean (λ) and the value to look up (x value).
4. Press [ENTER] .

**Binominal Distribution**

1. For $P(X = k)$, use the command `dbinom(k,n,p)`.
2. For $P(X \leq k)$, use the command `pbinom(k,n,p)`.

For example,

```
> dbinom(10, 30, .5)
[1] 0.0279816
```

Poisson Distribution

1. For $P(X = k)$, use the command `dpois(k, m)`.
2. For $P(X \leq k)$, use the command `ppois(k, m)`.

For example, the command below finds the probability of two or fewer events when they normally happen at a rate of three per unit.

```
> ppois(2, 3)
[1] 0.4231901
```

Common Continuous Distributions

Probability calculations for Normal random variables were already discussed in Chapter 1. Refer to that appendix if you need help. Probability calculations for uniform distributions are simply a case of identifying the base and height of the rectangle of interest. The calculation becomes the familiar $\text{Area} = \text{Probability} = (\text{base}) \times (\text{height})$. We detail how to perform calculations for exponential random variables below.



Exponential Distribution

1. Click an empty cell in the spreadsheet.
2. Click **Formulas** → **More Functions** → **Statistical**.
3. Scroll down the list of functions and double-click on the “EXPON. DIST” function.
4. In the “X” box, input the specified value. You will be calculating $P(X \leq \text{value})$.
5. In the “Lambda” box, input the value of λ , the mean occurrences per unit.
6. In the “Cumulative” box, enter **True**.
7. Click **Done** to find the results.



Exponential Distribution

1. Click **Applets** → **Distribution Calculator**.
2. Select “Exponential” for the distribution.
3. Enter the “Threshold”, which will be 0, and “1/Scale” (λ).
4. Use the radio buttons to select the type of calculation desired (input values to obtain a probability, or enter a probability to obtain a value).
5. Select the probability option (typically “ \leq ”) and enter the value k in the bottom right section and press the enter key to see shading on distribution and the result.



Exponential Distribution

1. Click **Calc** → **Probability Distributions** → **Exponential**.
2. Use the radio button to choose “Cumulative probability” if you wish to find $P(X \leq k)$ or “Inverse cumulative probability” if you want to find k given a probability.
3. Enter the “Scale” parameter (λ). Leave the “Threshold” at 0.
4. In most cases, you have a specific value k in mind; select “Input constant.” In the box next to this option, enter the specified value of k .
5. Click **OK** to find the results reported in the Session window.



Exponential Distribution

1. Click **Transform** → **Compute Variable**.
2. In “Target Variable,” enter a name for the new variable, then click in the “Numerical Expression” box.
3. In “Function Group” at right, select **CDF and Noncentral** for $P(X \leq k)$ and then **CDF.Exp**.
4. Enter the parameters: k , λ .
5. A new column will be added to the data set containing the desired probability. If necessary, use the Variable View tab to change the number of decimal places displayed.

CRUNCH IT!

CrunchIt! cannot calculate exponential probabilities.



TI-83/-84

TI calculators do not have a function to calculate probabilities for exponential distributions. You can, however, use numerical integration from the MATH menu (option 9). Enter the lower bound (0), the upper bound (k), the function $\lambda e^{-\lambda x}$, and the variable to integrate. An example is shown below. This finds $P(X < 1)$ when $\lambda = 2$. Use the arrows to move between the input boxes

A screenshot of a TI-83/84 calculator screen showing the integral function template: $\int_{\square}^{\square} (\square) d\square$. The input boxes are empty.

A screenshot of a TI-83/84 calculator screen showing the integral of $2e^{-2x}$ from 0 to 1. The result displayed is $.8646647168$.

**Exponential Distribution**

For $P(X \leq k)$, use the command `pexp(k,λ)`. For example,

```
> pexp(1,2)
[1] 0.8646647
```