1 Univariate random variables

October 12, 2009

1. Consider a random variable x that can take one of two values: zero and one, such that $P(1) = p$ and $P(0) = (1 - p)$ where $0 \leq p \leq 1$. Is x a continuous or a discrete random variable? discrete. Write down the density function for x

   $$P(x) \equiv f_X(x) \equiv f_X(x;p) = \begin{matrix} p^x(1-p)^{1-x} \text{ for } x = 0 \text{ or } 1 \\ 0 \text{ otherwise} \end{matrix} \quad \text{or}$$

   $$\begin{matrix} f(1) = p \\ f(0) = 1 - p \\ f(x) = 0 \text{ if } x \text{ does not equal } 0 \text{ or } 1 \end{matrix}$$

   Convince me that this is a legit density function. Since $0 < p < 1$, p and $(1 - p)$ are both positive, the function is never negative.

   $$f_X(0) + f_X(1) = (1 - p) + p = 1$$

   Does this density function have a name? It is the Bernoulli distribution.

   What is its $E[x]$?

   $$E[x] = \sum_{x=0}^{1} x f_X(x) = 0(1 - p) + 1(p) = p$$

2. What is its $var[x]$?

   $$\begin{aligned} var[x] &= E[(x - E[x])^2] = \sum_{x=0}^{1}(x - E[x])^2 f_X(x) \\ &= \sum_{x=0}^{1}(x - p)^2 f_X(x) = (0 - p)^2(1 - p) + (1 - p)^2 p \\ &= p^2(1 - p) + (1 - 2p + p^2)p = p^2 - p^3 + p - 2p^2 + p^3 \\ & \quad p(1 - p) \end{aligned}$$

3. Give me two examples of a random variable that would have a Bernoulli distribution: which side of a fair coin results from flipping a coin ($p = .5$), whether a queen is drawn when a card is randomly drawn from a deck ($p = \frac{4}{52}$)ir coin either drawing a queen from a deck.

4. What is $E[X]$? What is $var[Y]$?

5. For some density functions, determine $\Pr[(E[X]-\sigma) \leq x \leq (E[X]+\sigma)]$ and $\Pr[(E[X]-2\sigma) \leq x \leq (E[X]+2\sigma)]$. What do your examples suggest? Is this a general result that holds for all density functions. If so, can you demonstrate that it is a general result. Give some intuition and discuss the importance, if any, of what you determined.

   Explaining more, if one calculates the probability of the interval in assuming, for example, a normal distribution, the result will be a number that is independent of your choice of specific parameters values for the normal (the specific values you choose for the mean and the variance). For example, the probability of being one standard dev from the mean is always 60 some percent - I don't remember the exact number. Is this result true of all density functions. That is, is the result always the same number independent of what values you choose for the parameters in your chosen density function? And, if so, is the number the same number you would get for the normal?

6. Some of you told me that for a given family of density functions, $\Pr[(E[X]-\sigma) \leq x \leq (E[X]+\sigma)]$ is not a function of the parameter values the density takes. You concluded this on the basis of some examples where it was not a funciton of the parameters. Is your finding true for all density functions? Convince the reader that either it is or is not always true.

7. Specify some density function. Then determine $E(X) = \int\limits_{-\infty}^{\infty} xf(x)dx$ and present it graphically as an area. On your graph present both both $f(x)$ and $xf(x)$. Now include, in your presentation, the variance as a graphical area.

8. Consider a continuous random variable X with mean μ_X and finite variance σ_x^2. Someone has proven that $\Pr[|x-\mu_X| \geq r\sigma_X] \leq \frac{1}{r^2}$, which seems like it might imply some cool things. Provide me an example of this Chebyshev corollary in action, something cool but not too complicated.

9. Assume that the random variable Y is uniformly distributed on the interval $[a,b]$. Derive the probability density function of Y and its CDF. Find $E[Y]$ and $Var[Y]$.

10. If the variance of the random variable exists, show that $E[X^2] \geq (E[X])^2$.

11. What is a moment and what is a moment generating function, and why do we care?

12. Choose some density function, derive its moment generating function, and then use your moment generating function to derive three moments of your density function.

13. Made up a density starting by specifying the CDF. Convince the reader that yours is a legitimte CDF. Now derive the density function and some other properties of the distribution

14. If X is a continuous random variable with $var(X) = \sigma_x^2$, what is the variance of aX, where a is some parameter. Derive your result. What if $a = \frac{1}{n}$?

15. Can you prove that

$$\begin{aligned} var[X] &= E[(X - E[X])^2] \\ &= E[X^2] - (E[X])^2 \end{aligned}$$

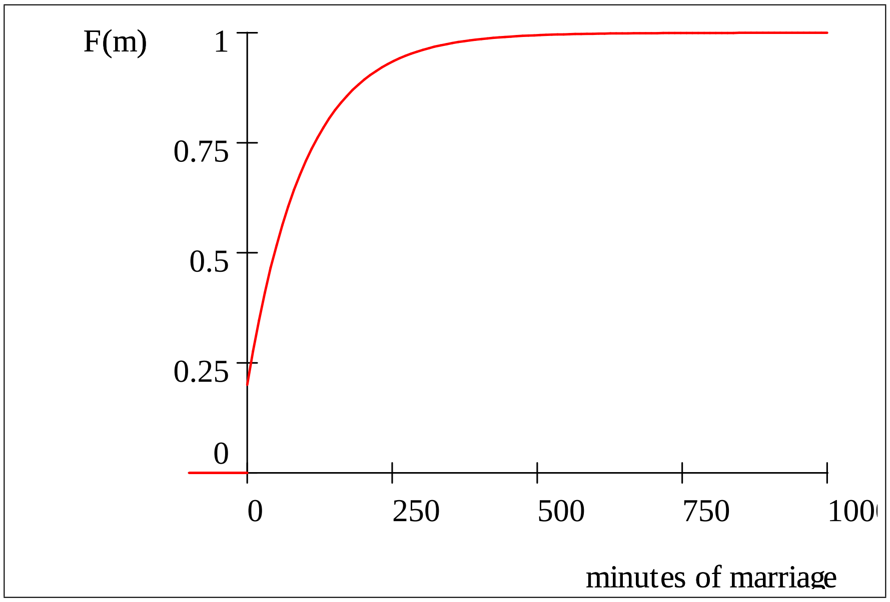
if $E[X^2]$ exists. If so, show a proof. If true, this formula gives us two ways to calculate a variance. This might come in handy. See MGB page 70, Theorem 4,

16. Consider the population of dead people who visited Las Vegas and assume we know how long each of them was married (in minutes) to their first spouse. Of course, some signficant proportion of these dead people were never married. Specify a function for the distribution of marriage time in this population (both $f_X(x)$ and $F_X(x)$, make sure they are consistent). Graph your distribution function and CDF. How would you describe the the distribution of this random variable? What is the expected marriage time for your population.

answer: I am going to change the names in the MGB example distribution function with respect to wait times at a stop sign (once you have legally come to a stop, sometimes there is no wait (you can go immediately), other times you wait). The MGB example is

$$F_M(m) = \begin{cases} 0 & if \quad m < 0 \\ (1 - pe^{-\lambda m}) & if \quad m \geq 0 \end{cases}$$

For example if $p = .8$ (20% of this population never marries) and $\lambda = .01$



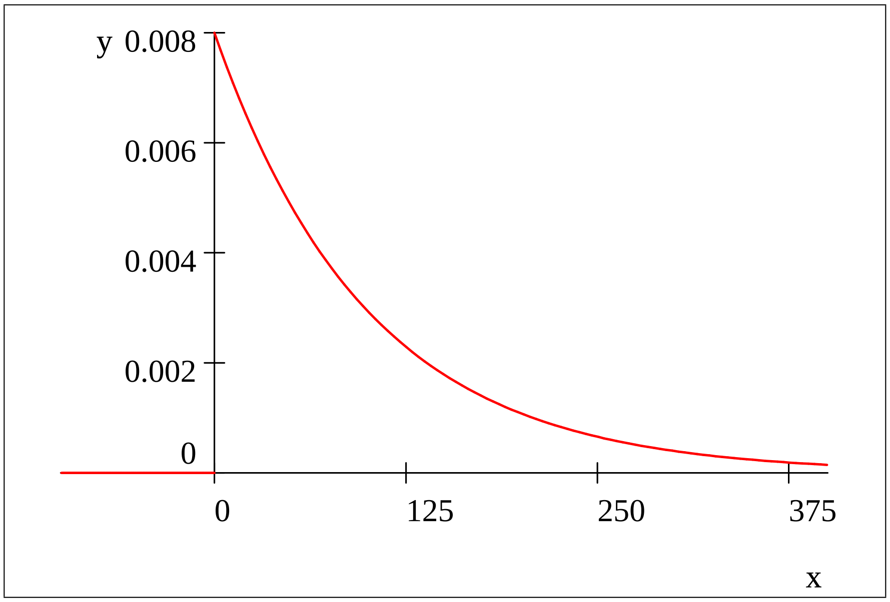
Marriages in this population are short. The question now is what is the corresponding distribution function. We know it is zero if $m < 0$, and $(1-p)$ if $m = 0$, so we just need to figure it out for positive m. The distribution function is the derivative of the CDF wrt m, so $\frac{\partial(1-pe^{-\lambda m})}{\partial(m)} = \lambda pe^{-\lambda m}$.[1] So, the distribution function is

$$f_M(m) = \begin{cases} 0 & if \quad m < 0 \\ (1-p) & if \quad m = 0 \\ \lambda pe^{-\lambda m} & if \quad m > 0 \end{cases}$$

Plotting this for the same values of p and λ,[2]

[1] $\frac{\partial(1-pe^{-\lambda m})}{\partial(m)} = \frac{\partial 1}{\partial m} - \frac{\partial(pe^{-\lambda m})}{\partial m} = -\frac{p\partial(e^{-\lambda m})}{\partial m}$
$= -pe^{-\lambda m}(-\lambda) = \lambda pe^{-\lambda m}$

[2] Note that $\int_0^{10000} 0.008\, e^{-0.01m} dm = 0.8$, so the total area under the distribution function sums to one.



I don't know why the spike of .2 at zero is not showing up.

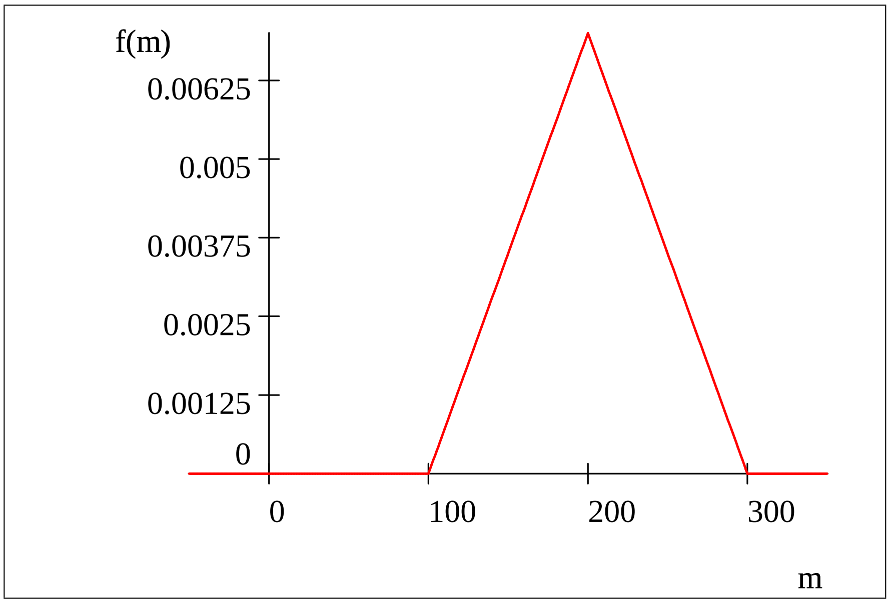
The expected number of marriages is, in general, $E[m] = \int\limits_0^\infty (1 - F_M(m))dm - \int\limits_{-\infty}^0 F_M(m))dm$. This formula for expected value works for both continuous and discrete distributions. In my example $\int\limits_{-\infty}^{<0} F_M(m))dm = 0$, so $E[m] = \int\limits_0^\infty (1 - F_M(m))dm = \int\limits_0^\infty (1 - (1 - pe^{-\lambda m}))dm = p\int_0^\infty e^{-m\lambda}\, dm$. So, if $p = .8$ and $\lambda = .01$, $E[m] = .8\int_0^\infty e^{-m(.01)}\, dm = 80.0$, the average marriage lasts 80 minutes. I got this formula for expected value from page 65 in my version of MGB. Another formula is $E[m] = \int\limits_{-\infty}^\infty m f_M(m)dm$. Which in this case simplifies to $E[m] = \int\limits_0^\infty m f_M(m)dm = .2(0) + \int\limits_0^\infty m(0.008\, e^{-0.01m})dm = 80.0$, which is the same answer -wow.

Now let's see if we can go come up with another example, but start with the distribution function. Let assume the probability of being married zero minutes is .30 (30%), that no one who gets married is married less than 100 minutes, and that

$$f_M(m.)\begin{cases} .000070(m.-100) & if \quad 100 \le m. \le 200 \\ .014-.000070(m.-100) & if \quad 300 \ge m. > 200 \\ 0 & if \quad m. > 300 \end{cases}$$

So,

$$f_M(m.) = \begin{cases} 0 & if \quad m. < 0 \\ .3 & if \quad m. = 0 \\ 0 & if \quad 0 < m. < 100 \\ .000070(m.-100) & if \quad 100 \le m. \le 200 \\ .014-.000070(m.-100) & if \quad 300 \ge m. > 200 \\ 0 & if \quad m. > 300 \end{cases}$$



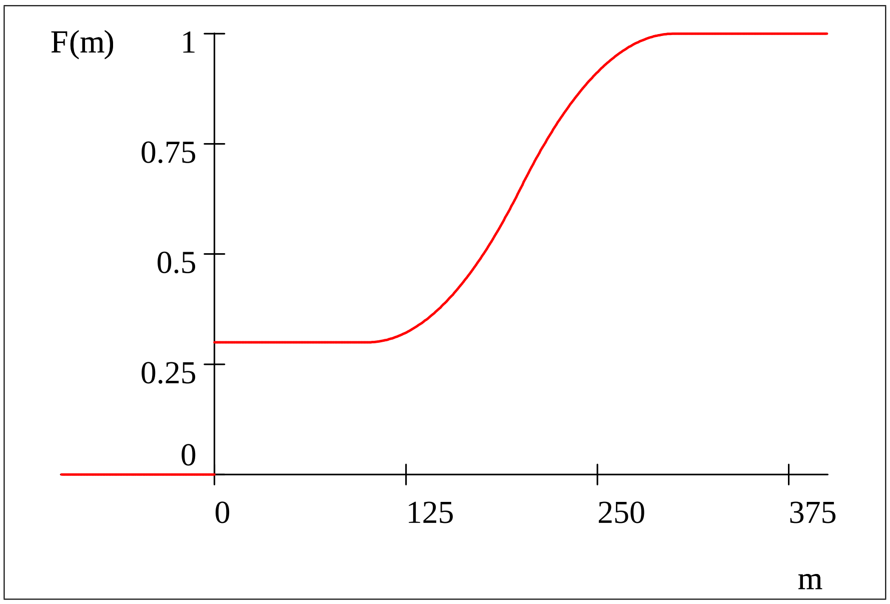
I don't know what I am doing wrong, but I can't make the spike of .3 show up at zero. Let's see what the the corresponding CDF looks like.

$$F_X(m) = \begin{cases} 0 & if \quad m. < 0 \\ .3 & if \quad 0 \le m. < 100 \\ .65 + 3.5 \times 10^{-5} m.^2 - 0.007\, m. & if \quad 100 \le m. \le 200 \\ -2.15 + 0.021\, m. - 3.5 \times 10^{-5} m.^2 & if \quad 300 \ge m. > 200 \\ 1 & if \quad m. > 300 \end{cases}$$

Note that in defense of the above, $\int(.000070(\ m.-100))d\ m. = 3.5\times10^{-5}$ $m.^2 - 0.007\ \ m.$ and $\int_{100}^{200}(.000070(\ m.-100))d\ m = 0.35$,

$\int(.014-.000070(m-100))dm = 0.021\,m-3.5\times10^{-5}\,m^2$ and $\int_{200}^{300}(.014-.000070(m-100))dm = 0.35$. The constants of integration were chosen so that $F_M(100) = .30$ and $F_M(100) = .65$.[3]

Graphing the CDF



some comments on student answers: Your distribution must have a spike at zero minutes; that is, your density function must have a discrete component. So, it must be either a discrete density function (with many spikes) or a discrete/continuous density function (a mixture of densities). Said another way, if your density function is continuous, the probability of marriage time being any specific amount (including zero) is zero, which violates the condition that a proportion of the population never marries.

a caveat: Some groups got around the necessity of a spike at zero by redefining marriage. These groups effectively said that a marriage was not a marriage if is lasted less than a certain amount of time. In other words, rather than specifying a discrete/continuous distribution, you redefined marriage to the distribution had no discrete component. Inventive.

17. MGB says that the random variable X is actually a function $X()$. What are they talking about?

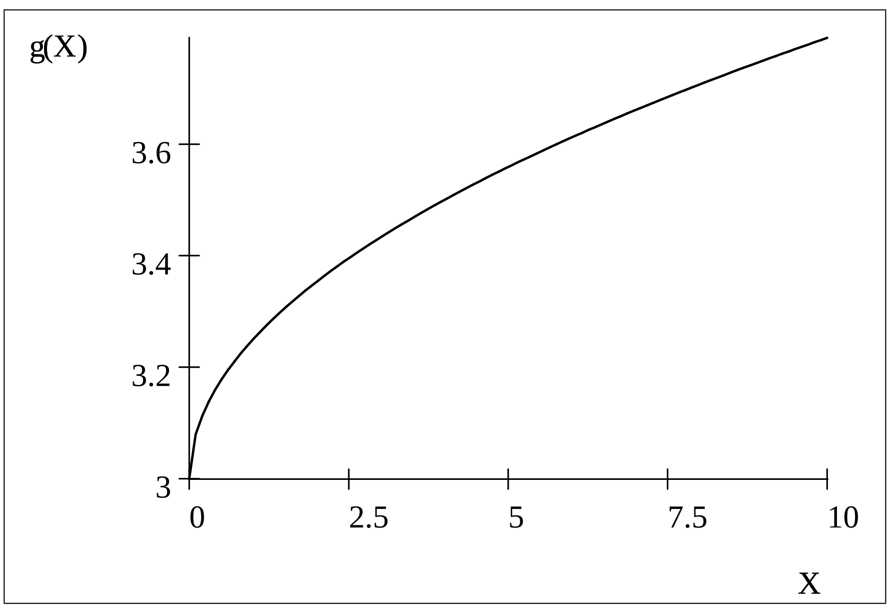
    answer: Experiments have outcomes and events. We want to associate a unique number with each outcome, and assocate a unique number with each event of interest. For example, if the experiment has N_Ω outcomes,

[3] $0.021\,(200) - 3.5\times10^{-5}(200)^2 = 2.8$ but it need so be .65, so substract 2.15. $3.5\times10^{-5}(100)^2 - 0.007(100)\,) = -0.35$. Needs to be .3 so add 0.65.

our random variable representing outcomes has to take N_Ω different values. This is a mapping from outcomes to points on the real line: a function. If the experiment has an infinite number of outcomes, the rv representing those outcomes needs to take an infinite number of values, one for each outcome. Note that this mapping is sometimes so obvious we don't even notice it. For example, if the outcome of an experiment is an individual's weight, the the outcome is typically respresented as a number representing the individual's weight (for example, in pounds, kilos, or tons) - the mapping is very direct. Alternatively, the outcome might be a head or a tail, or a male or a female. Then we map heads into one number and tails into another. The choice of the two numbers is arbitrary, so the mapping is not unique. Typically we use a zero and one, but we could use -9.2 and 3.57. If the outcomes are black male, black female, white male, and white female, we represent this with a rv that can take only four values - this is a functional mapping.

18. Convince me, or not, that $E[3+.25x^{.5}] \geq 3+.25(E[X])^{.5}$

    answer: Consider the Jensen Inequality: $E[g(x)] \geq g(E[X])$ if $g()$ is a convex function. So, if $g(X) = 3+.25x^{.5}$ is a convex function $E[3+.25x^{.5}] \geq 3+.25(E[X])^{.5}$. Is it?



19. So, average height in the U.S. population is 68 inches ($E[H]=68$). Consider the statement by Senator Snerd who is running for President,"There is a 98% chance that a randomly-selected individual is at least 72 inches tall." Discuss.

answer: Consider the Chebyshev Inequality: $P[g(X) \geq k] \leq \frac{E[g(X)]}{k}$ for all $k > 0$. A special case is $P[X \geq k] \leq \frac{E[X]}{k}$. So if $E[X] = 68$ and $k = 72$, then $P[X \geq 72] \leq \frac{68}{72} = 0.944\,44$. That is the probability of drawing someone who is at leat 72 inches tall has to be less than .94, not matter how height is distributed, so it cannot be .98. The above statement by Senator Snerd has to be false. MGB say the inequality holds whether X is continuous or discrete, but only provides a proof for the case where X is continous. In a corollary, MGB had the assumption that X has finitie variance.

Another correct answer would proceed along the following lines: Demonstrate that for Snerd to be correct requires some heights in the population to be negative, and since people cannot have negative height, Snerd must be wrong.

One could answer this question assuming X has a normal distribution (why it cannot), but this would be greatly simplifying the question.

20. Consider a continuous random variable with density $f_X(x)$. Describe in word $f_X(4)$. Now consider a discrete random variable with probability density function $f_X(x)$ and describe in words $f_X(4)$.

    answer: For the discrete rv X, $f_X(4)$ is the probability that X takes the value 4; it with be a non-negative number less than one. On the other hand, if X is a continuous random variable, $f_X(4)$ is simply the height of the density function at $X = 4$. It is a non-negative number that is not bounded from above by one (it can be greater than one). It is not the probability that $X = 4M$. The probability that X is exactly 4 is zero.

21. Consider the population of people named Fred who had difficult childhoods. Assume desire, D, for chocolate in this population is normally distributed with mean μ and variance σ^2. D_i is individual i's desire for chocolate. Let C_i be the number of chocolate bars individual i consumes. Assume $C_i = 0$ if $D_i \leq 0$ and $C_i = D_i$ if $D_i > 0$. Derive and express the density function for chocolate-bar consumption in this population. Enhance your presentation with graphs. Is there a name for the type of distribution you created? And, if so, what is the name. Relate the type of distribution you created in this question with the type that described marriage Las Vegas style. Now assume

$$f_D(d) \left\{ \begin{array}{ccc} 0 & if & d < 0 \\ \lambda e^{-\lambda d} & if & d \geq 0 \end{array} \right.$$

    and that $C_i = 0$ if $D_i \leq \alpha$, $\alpha > 0$ and $C_i = D_i$ if $D_i > \alpha$. Assume $\lambda > 0$. For this case, derive the density function for chocolate-bar consumption.

    answer: what we are doing here is deriving a density function by *censoring* some existing density function. For more details, see the student notebooks on the course web page. Green (third edition, pages 959-961) discusses censoring, in particular the censored normal distribution. Tobit models were developed to deal with censored data.

22. Consider the population of people named Fred who had difficult childhoods. Assume desire, D, for chocolate in this population is normally distributed with mean μ and variance σ^2. D_i is individual i's desire for chocolate. Let C_i be the number of chocolate bars individual i consumes. Assume $C_i = D_i$ if $D_i \geq 0$. Now assume everyone with negative desire is shot dead - not liking chocolate is a sin. Derive and express the density function for chocolate-bar consumption in the remaining population. Enhance your presentation with graphs. Is there a name for the type of distribution you created? And, if so, what is the name. Relate the type of distribution you created in this question with the type that described marriage Las Vegas style. Now assume

$$f_D(d) \begin{cases} 0 & if \quad d < 0 \\ \lambda e^{-\lambda d} & if \quad d \geq 0 \end{cases}$$

where $\lambda > 0$. Assume $C_i = D_i$ if $D_i \geq \alpha$. Assume $\alpha > 0$. Now assume everyone with desire less than α is shot dead. Derive and express the density function for chocolate-bar consumption in the remaining population. Enhance your presentation with graphs.

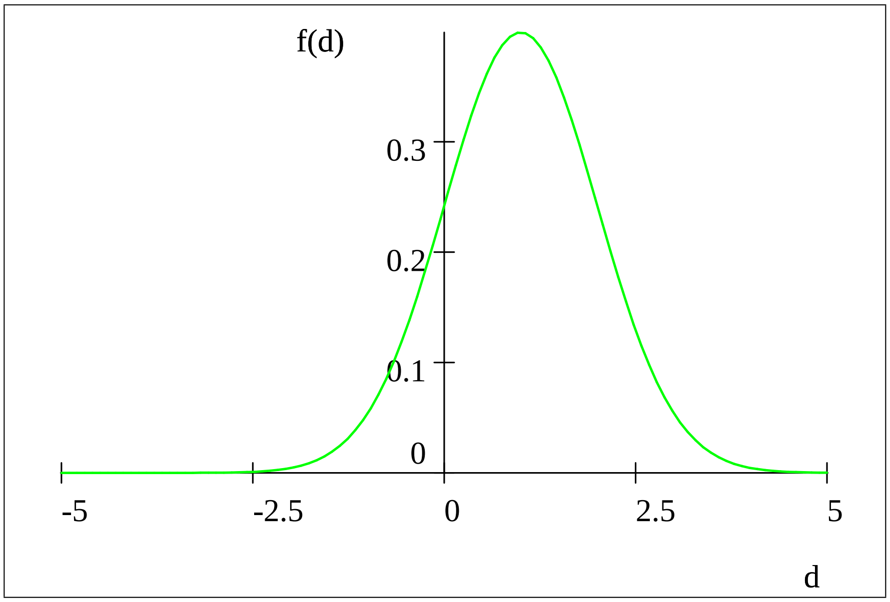
answer: The density function for chocolate bars, C, is a truncated-normal distribution, truncated on the left at $c = 0$. Let $\phi_{\mu,\sigma}$ denote the density function for the normal distribution with mean μ and variance σ^2, and let $\Phi_{\mu,\sigma}$ denote its correspond CDF. Specifically the density function for C is

$$f_C(c) = \begin{cases} 0 & if \quad c \leq 0 \\ \frac{\phi_{\mu,\sigma}(c)}{1-\Phi_{\mu,\sigma}(0)} & if \quad c > 0 \end{cases}$$

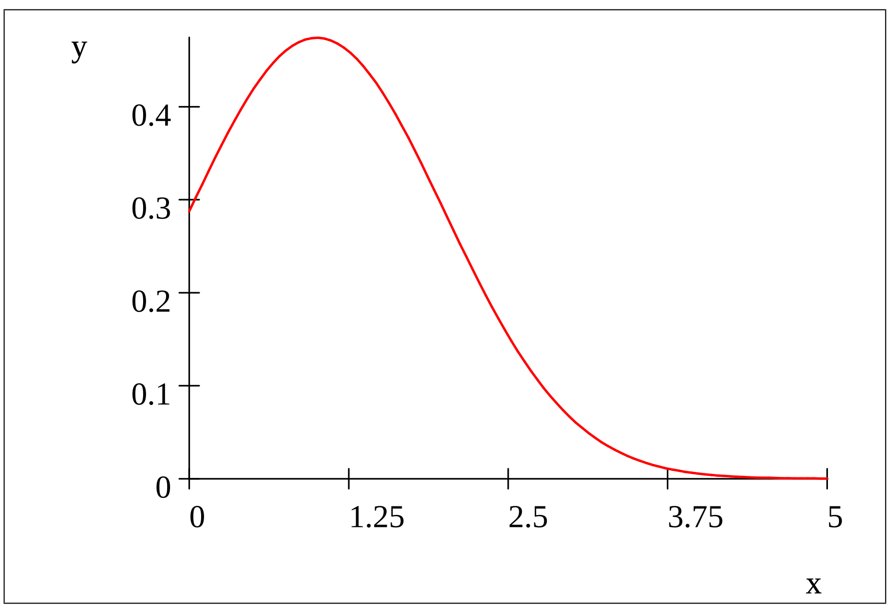
where $\frac{\phi_{\mu,\sigma}(c)}{1-\Phi_{\mu,\sigma}(0)} = \frac{\frac{1}{\sigma\sqrt{2\pi}} e^{-(c-\mu)^2/2\sigma^2}}{1-\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{0} e^{-(m-\mu)^2/2\sigma^2} dm}$. In explanation, we lop off the normal to the left of zero (make the density there zero). But then, the area under what remains is less than one, so not a density function. To get the area to be one, we need to divide by the area under what is left, which is $1 - \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{0} e^{-(m-\mu)^2/2\sigma^2} dm$. For more details see sec 4.3 Contagious Distributions and Truncated Distributions in MGB. Green (third edition, page 949-952) discusses the truncated normal in detail.[4]

For example, if one assumes $\mu = 1$ and $\sigma = 1$, the density function for desire, which is normally distributed, is

[4] In terms of notation and terminology, if the density function is $f(x)$, then its truncated form, truncated at α, is written $f(x|x > \alpha)$.

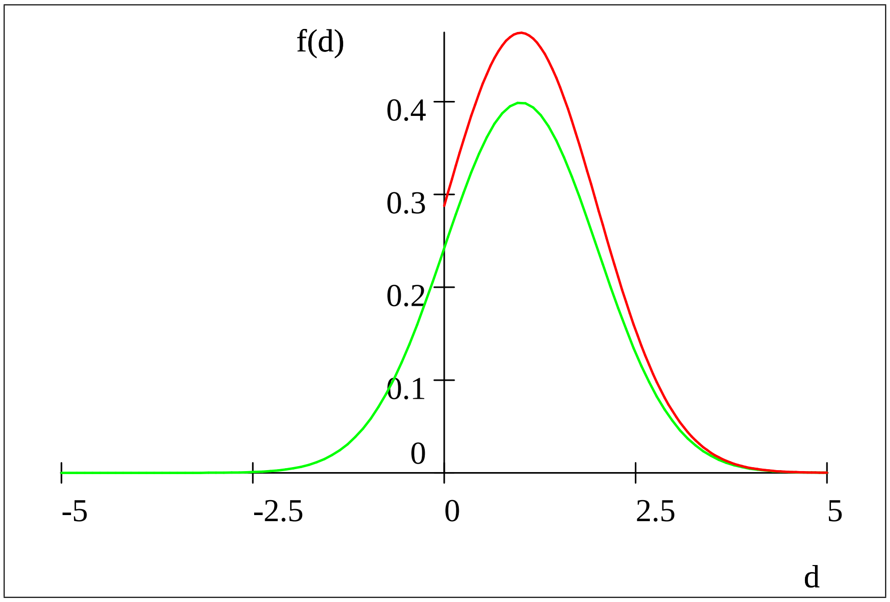


And $1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-(m-1)^2/2} dm = 0.841\,34$, so we are lopping off approximately 16% of the distribution (they get shot). So, the truncated density function for chocolate bars, in the positive range, is $f_C(c) = \frac{\frac{1}{\sqrt{2\pi}} e^{-(c-1)^2/2}}{0.841\,34}$.[5]



[5] Checking to make sure this is a density function. $\int_0^\infty \left(\frac{\frac{1}{\sqrt{2\pi}} e^{-(c-1)^2/2}}{0.841\,34} \right) dc = 1.0$

Comparing the two densities side by side



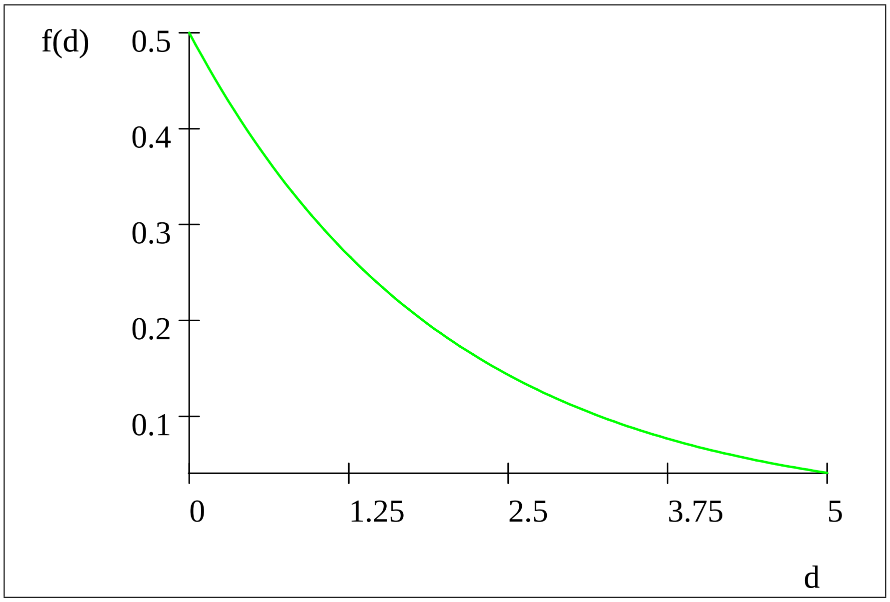
Note that the mean of the truncated normal is greater than the mean of the untruncated normal (this we can "see" from the graph). It is also possible to prove that the variance of the truncated normal is less than the variance of the normal.

This derived distribution for chocolate bars was not derived in the same way as the distribution of marriage times for dead people who had visited Las Vegas. In both cases, the density to the left of some point was set to zero. Here it is "added back in" by shifting upward the remaining positive density. For marriage times the part of the density that was set to zero was converted into a spike.

Now consider the alternative density function

$$f_D(d) \begin{cases} 0 & if \quad d < 0 \\ \lambda e^{-\lambda d} & if \quad d \geq 0 \end{cases}$$

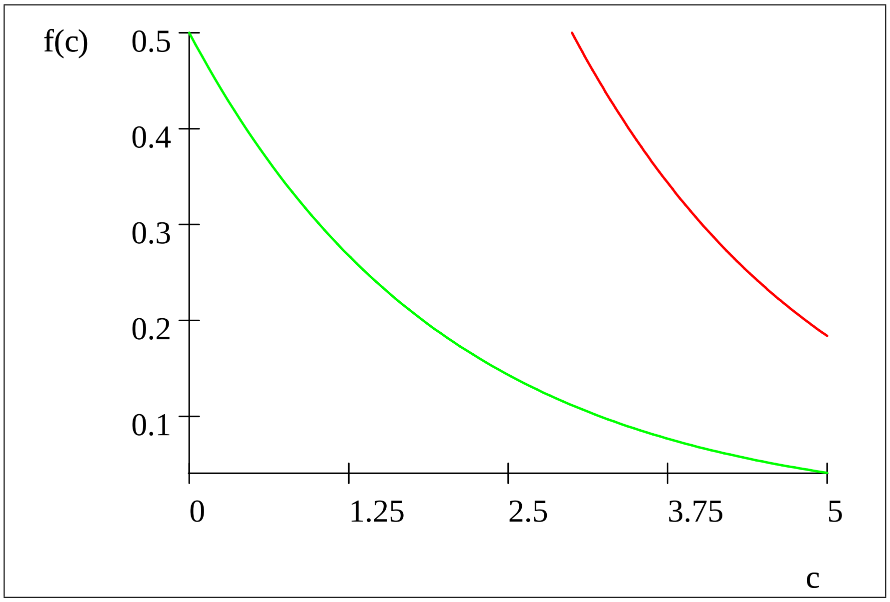
This is the Exponential density function. It has positive mass from zero to infinity. If, for example, $\lambda = .5$ it looks as follows



So, in this case

$$f_C(c) = \begin{cases} 0 & if \quad c \leq \alpha \\ \frac{\lambda e^{-\lambda c}}{1-\int_0^{\alpha}(\lambda e^{-\lambda d})dd} & if \quad c > \alpha \end{cases}$$

For example, if $\lambda = .5$ and $\alpha = 3$, $1 - \int_0^3 (.5e^{-.5d})dd = 0.223\,13$. So, approximately 78% of the positive density is set to zero. The following, in red, is a graph of this truncated Exponential. $\frac{.5e^{-.5c}}{0.223\,13}$

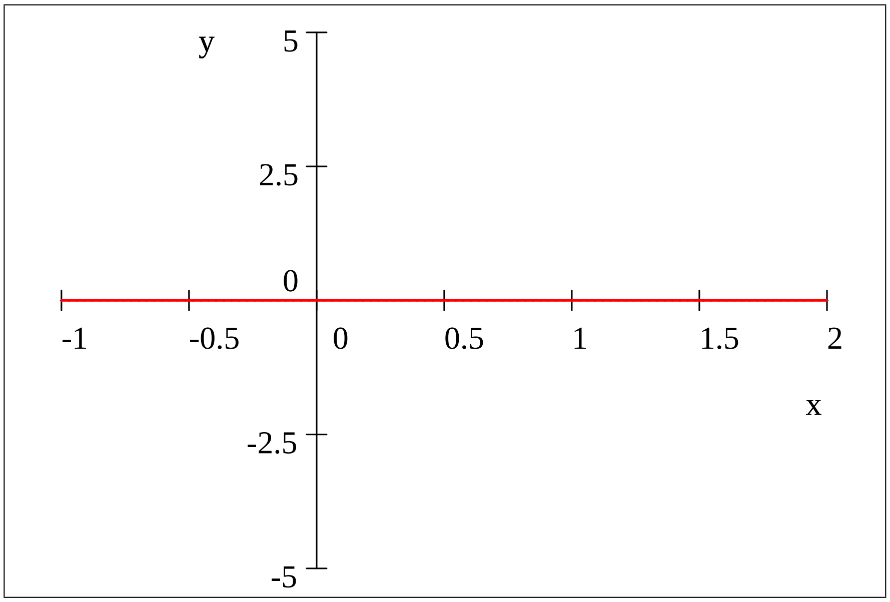


23. Assume the rv X is normally distributed with mean μ and variance σ^2. Creat a new distribution by cutting off some both tails of this normal distribution. Explain what you are doing with words and graphs. Convince me your new density function is in fact a density function. Is there a name for this type of distribution?

24. Consider the experiment of the birth of a baby where the probability of the birth of a girl is $p = .5$ and the probability of the birth of a boy is .5. Let the random variable X denote the number of girls. Specify the density function and CDF for X. Graph both functions

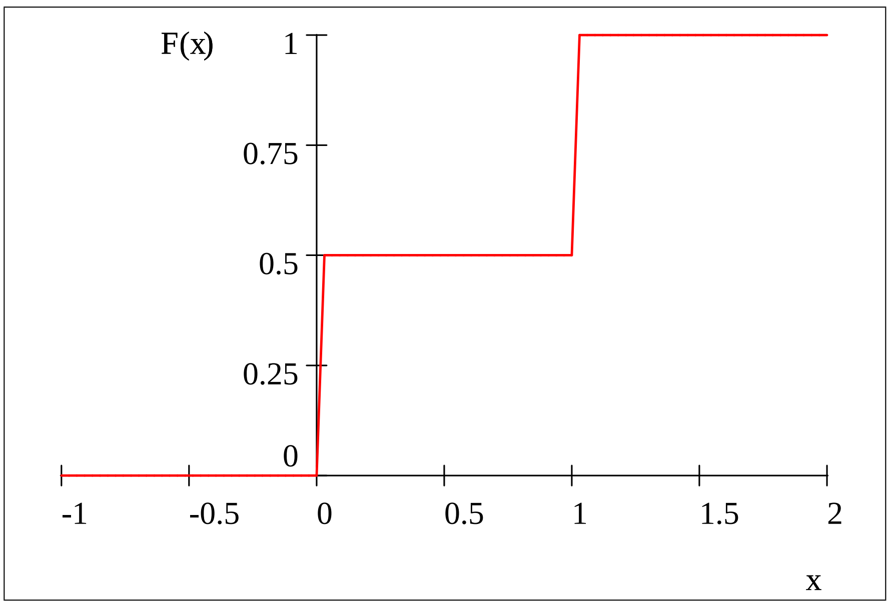
    answer:

$$f_X(x) = \begin{cases} 0 & if \quad x < 0 \\ .5 & if \quad x = 0 \\ 0 & if \quad 0 < x < 1 \\ .5 & if \quad x = 1 \\ 0 & if \quad x > 1 \end{cases}$$



and

$$F_X(x) = \begin{cases} 0 & if \quad x < 0 \\ .5 & if \quad 0 \leq x < 1 \\ 1 & if \quad x \geq 1 \end{cases}$$

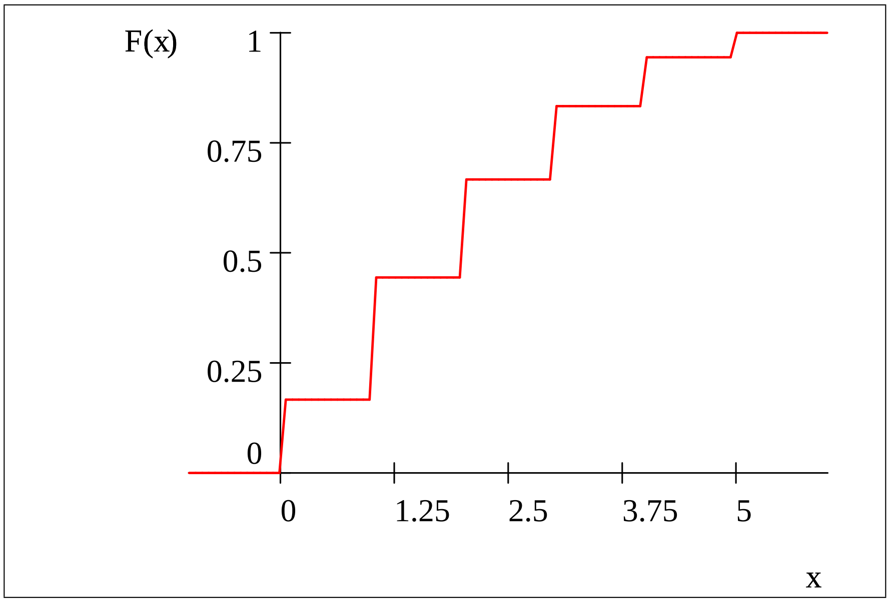


25. Consider the experiment of tossing two fair dice. Indentify the outcome space, Ω. Let the rv variable X denote the absolute difference between

the face value of the first and second die (e.g. 1,2, ..., 5). Specify the density function and CDF of this rv and graph both the density function and the CDF.

answer: There are 36 possible outcome. Consider a graph of dots (each dot representing one of the 36 outcomes) where the horizontal axis is the number on the first die (1 to 6) and the vertical axis is the number of die on the second die. The rest is counting. There are six ways the absolute difference is zero $((1,1),(2,2),(3,3),(4,4),(5,5)$ and $(6,6))$There are ten ways the absolute difference is one $((1,2),\ (2,1),(2,3),(3,2),(3,4),(4,3),(4,5),(5,4),(5,6)$ and $(6,5))$. Eight ways the absolute difference is two $((1,3),(2,4),(3,1),(3,5),(4,2),(4,6),(5,3),(6,4))$. Six ways the absolute difference is three $((1,4),(2,5),(3,6),(4,1),(5,2),(6,1))$. Four ways the absolute difference is four $((1,5),(2,6),(5,1),(6,2))$. And, two way the absolute difference is five $((1,6),(6,1))$. So the CDF is

$$F_X(x) = \begin{cases} 0 & if \quad x < 0 \\ \frac{6}{36} & if \quad 0 \le x < 1 \\ \frac{16}{36} & if \quad 1 \le x < 2 \\ \frac{24}{36} & if \quad 2 \le x < 3 \\ \frac{30}{36} & if \quad 3 \le x < 4 \\ \frac{34}{36} & if \quad 4 \le x < 5 \\ \frac{36}{36} & if \quad x \ge 5 \end{cases}$$



and the density function, $f_X(x)$ is

26. Consider a zoo that has only chimps and elephants where $poop_j$ is the weight (in pounds) of *poop* animal j produces on a given day, $S_j = 1$ if

the animal is a elephant, and zero otherwise. The amount of *poop* animal j produces on a given day is a draw from a Poisson distribution (some animals on some days are constipated) with parameter λ_j where $\lambda_j = \alpha + \beta S_j$. The director of the zoo likes only chimps, but knows that chimps like to ride on the backs of elephants, so has 90 chimps and 10 elephants (lots of chimps can ride one elephant). Write down the density function for *poop* for this population. Explain why this is the density function. Derive the expected amount of *poop* produced per day. Comment briefly on the likelihood of the *poop* being Poisson distributed rather than normally distributed.

answer: The distribution function $f(poop)$ is what is called a mixture distribution or a contagious distribution. In this case, a mixture of two densities: one for chimps and the one for elephants. For elephants, the distribution function is $f_e(poop) = \frac{e^{-(\alpha+\beta)}(\alpha+\beta)}{poop!}$ for $x = 0, 1, 2, ...$, and 0 otherwise. For chimps, it is $f_c(poop) = \frac{e^{-\alpha}\alpha}{poop!}$. At this zoo, chimps appear with probability $p_c = .9$ and elephants appear with probability $p_e = .1$. So $f(poop) = .1(\frac{e^{-(\alpha+\beta)}(\alpha+\beta)}{p!}) + .9(\frac{e^{-\alpha}\alpha}{p!})$ for $x = 0, 1, 2, ...$, and 0 otherwise. This is an application of the following theorem. If one has a sequence $f_o(x)f_1(x), ...$ of density functions and a sequence $p_o, p_1, ...$ such that $p_i \geq 0$ and $\sum p_i = 1$, then $f(x) = \sum p_i f_i(x)$.

The expected amount of *poop* is $.9\alpha + .1(\alpha + \beta) = 1.0\alpha + 0.1\beta$, the simple weighted average.

On a given day, an animal can't poop a negative amount. The Poisson is consistent with this restriction, the normal is inconsistent with this restriction: the normal assocates positive probability with negative amounts of poop. The Poisson assumes poop is discretely distributed (animals poop in units of $0, 1, 2, ..$ but not 1.34 units). The normal allows all animals to poop in other than integer amounts. The latter is more likely. On the other hand, the Poisson allows for the possibility that on a give day the animal produces no poop, the normal does not.

27. In the class notes, we modeled number of times married using the Poisson distribution - a discrete distribution of non-negative intergers. Alternatively, we could have modled number of times married with a Geometric distribution or a Negative binomial distribution. These are also discrete distributions of non-negative integers. How would you choose which was the most appropriate density for this random variable?

28. Common measures of central tendency are the mean the median and the mode. Consider now a forth measure. To define this forth measure, we need to introduce the concept of a *quantile*. MGB define the th q^{th} quantile of the random variable X as the smallest number ξ satisfying $F_X(\xi) \geq q$. They denote this number ξ_q. For example, $\xi_{.5}$ is the $.5^{th}$ quantile, also known as the median. In contrast, $\xi_{.25}$ is the $.25^{th}$ quantile: 25 percent of

the density lies below $\xi_{.25}$. So, now that we know what a quantile is, define a forth measure of central tendency as $.5(\xi_{.25} + \xi_{.75})$. Specify a density density where these four measures of central tendency are all different - ideally quite different.