

Chapter 10

Continuous Random variables

We will skip over the details of the distinction between discrete and continuous variables, and the connection between them. The distinction is rather artificial since the variables continuous/discrete can be defined in terms of Borel sets and the probability as a *Lebesgue measure* on these sets. This is covered more rigorously in *measure theory*. Instead we will *rederive* everything!

10.1 Probability density

Suppose X is a random variable, within the domain $a \leq x \leq b$. Then we define the *probability density function* $f_X(x)$ as follows:

$$\boxed{P(x \leq X \leq x + dx) = f_X(x)dx} \quad (10.1)$$

The (cumulative) *probability distribution function* is defined as the integral:

$$F_X(x) = P(a \leq X \leq x) = \int_a^x f_X(x')dx' \quad (10.2)$$

Then, in general:

$$P(c \leq X \leq d) = F_X(d) - F_X(c) \quad (10.3)$$

Since the total probability should add up to 1, we have:

$$P(a \leq X \leq b) = \int_a^b f_X(x')dx' = 1 \quad (10.4)$$

The *moments* of the distribution are defined as in the discrete case (replace $\sum_i \rightarrow \int_a^b dx$).

$$\boxed{\mathbb{E}(X) = \mu = \int_a^b x f_X(x)dx} \quad (10.5)$$

Then it follows that:

$$\mathbb{E}(X^2) = \int_a^b x^2 f_X(x)dx \quad (10.6)$$

and one obtains:

$$\boxed{\text{var}(X) \equiv \mathbb{E}(X^2) - (\mathbb{E}(X))^2} \quad (10.7)$$

The moment-generating function can be defined similarly,

$$M_X(t) \equiv \int_{-\infty}^{+\infty} e^{tx} f_X(x)dx \quad (10.8)$$

And finally, since the distribution is the integral of the density, the derivative of the distribution is the probability density:

$$\frac{d}{dx} F_X(x) = \frac{d}{dx} \int_a^x f_X(x')dx' = f_X(x) \quad (10.9)$$

10.1.1 Example

Suppose the probability has a *uniform* density, over the range, $a \leq x \leq b$. That is:

$$f_X(x) = \frac{1}{b-a} \quad , \quad a \leq x \leq b \quad . \quad (10.10)$$

and $f_X(x) = 0$ everywhere else. Then, for $a \leq x \leq b$, we have:

$$F_X(x) = \int_a^x f_X(x)dx = \int_a^x \frac{1}{(b-a)}dx = \frac{(x-a)}{(b-a)} \quad . \quad (10.11)$$

Furthermore, it follows that:

$$\mu = \mathbb{E}(X) = \int_a^b x f_X(x)dx = \int_a^b x \frac{1}{(b-a)}dx \quad (10.12)$$

$$\mathbb{E}(X) = \frac{1}{(b-a)} \left[\frac{1}{2}x^2 \right]_a^b = \frac{1}{(b-a)} \frac{1}{2}(b^2 - a^2) = \frac{1}{2}(b+a) \quad (10.13)$$

The second moment is given by the expression:

$$\mathbb{E}(X^2) = \int_a^b x^2 \frac{1}{(b-a)}dx = \frac{1}{(b-a)} \frac{1}{3}(b^3 - a^3) = \frac{1}{3}(a^2 + ab + b^2) \quad . \quad (10.14)$$

Then the variance is given by:

$$\text{var}(X) = \frac{1}{3}(a^2 + ab + b^2) - \frac{1}{4}(b+a)^2 = \frac{1}{12}(b-a)^2 \quad . \quad (10.15)$$

10.1.2 Example

The *exponential distribution* is defined by the probability density:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (10.16)$$

(a) Calculate $\int_{-\infty}^{\infty} f_X(x) dx$, (b) $\mathbb{E}(X)$ and (c) $\text{var}(X)$.

10.1.3 Answer

(a)

$$\int_{-\infty}^{\infty} f_X(x) dx = \int_0^{\infty} \lambda e^{-\lambda x} dx = [-e^{-\lambda x}]_0^{\infty} = 1 \quad (10.17)$$

(b)

$$\mathbb{E}(X) = \int_0^{\infty} x f_X(x)dx = \int_0^{\infty} x \lambda e^{-\lambda x} dx \quad (10.18)$$

Integrating by parts:

$$\int u \frac{dv}{dx} dx = [uv] - \int v \frac{du}{dx} dx \quad (10.19)$$

$$\int_0^{\infty} x \lambda e^{-\lambda x} dx = [x(-e^{-\lambda x})]_0^{\infty} - \int_0^{\infty} (-e^{-\lambda x}) 1 dx = 0 + \int_0^{\infty} e^{-\lambda x} dx = \frac{1}{\lambda} \quad (10.20)$$

$$\mathbb{E}(X^2) = \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx \quad (10.21)$$

This can be calculated by integration by parts twice, however, there is a more efficient alternative - differentiating twice.

$$\mathbb{E}(X^2) = \lambda \int_0^{\infty} x^2 e^{-\lambda x} dx = \lambda \left(\frac{\partial}{\partial \lambda} \right)^2 \int_0^{\infty} e^{-\lambda x} dx \quad (10.22)$$

Then,

$$\mathbb{E}(X^2) = \lambda \left(\frac{\partial}{\partial \lambda} \right)^2 \left[\frac{1}{\lambda} \right] = \frac{2}{\lambda^2} \quad (10.23)$$

$$\text{var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \left(\frac{2}{\lambda^2} \right) - \left(\frac{1}{\lambda} \right)^2 = \frac{1}{\lambda^2}. \quad (10.24)$$

10.1.4 Example

The *normal* or *Gaussian* distribution with parameters, μ, σ is given by the density:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right), \quad -\infty < x < +\infty. \quad (10.25)$$

Then (see homework), we can show that:

$$\mathbb{E}(X) = \mu \quad \text{and} \quad \text{var}(X) = \sigma^2. \quad (10.26)$$

This distribution is denoted by the symbol $N(\mu, \sigma^2)$. Values of the *standard* normal distribution $N(0, 1) \Rightarrow$ mean = 0, variance = 1, are widely tabulated. The density in this case is given the special notation $\phi(x)$:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}. \quad (10.27)$$

The corresponding probability distribution is defined as follows:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{1}{2}x^2} dx = P(Z \leq z). \quad (10.28)$$

10.2 The Markov and Chebyshev Inequalities

Consider a continuous random variable X , such that X is non-negative: $X \geq 0$. This is not unduly restrictive since we can always add a value onto the variable to ensure this is the case. So if b is the (negative) lower limit of X , then the variable, $Y = X + b \geq 0$. Then, denoting the mean by $\mathbb{E}(X) = \mu$, the following inequality is true for any $a > 0$

$$\boxed{P(X \geq a) \leq \frac{\mu}{a}}. \quad (10.29)$$

This is called the *Markov inequality*.

10.2.1 Proof

Firstly, consider the left-hand side expression, which is by its definition the integral:

$$P(X \geq a) \equiv \int_a^{\infty} f(x) dx. \quad (10.30)$$

Then, we can divide the range of integration for the mean as follows:

$$\mu = \int_0^{\infty} x f_X(x) dx = \int_0^a x f_X(x) dx + \int_a^{\infty} x f(x) dx. \quad (10.31)$$

This leads to the expression:

$$\mu - \int_a^{\infty} x f_X(x) dx = \int_0^a x f_X(x) dx, \quad (10.32)$$

and since the right-hand-side is clearly ≥ 0 , since the integrand is always ≥ 0 , we have:

$$\mu - \int_a^\infty (x - a + a)f_X(x) \geq 0 \quad , \quad (10.33)$$

in which a has been added and subtracted to a factor in the integrand. Now separating these two terms gives the result:

$$\mu - \int_a^\infty (x - a)f_X(x)dx - a \int_a^\infty f_X(x)dx \geq 0 \quad . \quad (10.34)$$

Thus:

$$\mu - aP(X \geq a) \geq \int_a^\infty (x - a)f(x)dx \geq 0 \quad (10.35)$$

Finally giving the Markov inequality :

$$P(X \geq a) \leq \frac{\mu}{a} \quad . \quad (10.36)$$

10.3 Chebyshev Inequality

The *Chebyshev inequality* takes the form:

$$\boxed{P(|X| \geq a) \leq \frac{\mathbb{E}(X^2)}{a^2}} \quad . \quad (10.37)$$

10.3.1 Proof

Note firstly that $|X| \geq a \Leftrightarrow X^2 \geq a^2$, therefore,

$$P(|X| \geq a) = P(X^2 \geq a^2) \quad (10.38)$$

Now using the Markov inequality for the variable X , we can write

$$P(|X| \geq a) = P(X^2 \geq a^2) \leq \frac{\mathbb{E}(X^2)}{a^2} \quad . \quad (10.39)$$

Thus giving the Chebyshev inequality. There is a simple extension of this inequality which follows from a change of variable. Replacing X by $X - \mu$, and recognising that $\mathbb{E}((X - \mu)^2) \equiv \sigma^2$, leads to the result:

$$\boxed{P(|X - \mu| \geq a) \leq \frac{\sigma^2}{a^2}} \quad . \quad (10.40)$$

which is the more common form quoted for the inequality.

10.4 Change of Measure

Recall that the expected value of any function of a continuous random variable can be written as:

$$\mathbb{E}(u(X)) = \int_{-\infty}^{+\infty} u(x)f_X(x)dx \quad . \quad (10.41)$$

Since, by definition of the density and distribution we have:

$$dF_X(x) = f_X(x)dx \quad .$$

Then changing the variable of integration to the probability distribution rather than the random variable, the expected value can be written as:

$$\boxed{\mathbb{E}(u(X)) = \int_0^1 u(x)dF_X(x)} \quad . \quad (10.42)$$

This is an expression of the Radon-Nikodym theorem, which changes variable from x to F . Suppose, although we are interested in the same random process, we wanted to change the random variable to another random variable. This new variable, call it y , would have a different measure, which we call G .

In general the variable of integration can be changed so long as there is a one-to-one correspondence between the new variable and the old variable. That is, a unique inverse function exists. The change of independent variable will then lead to a change of probability density (in the new variable). Often we speak of a *change of measure* when the probability distribution is modified in this way.

Suppose we define a new random variable, Y , such that this inversion is possible:

$$Y = h(X) \quad , \quad \text{and} \quad X = h^{-1}(Y) \quad . \quad (10.43)$$

where $h(X)$ is *differentiable*. Then:

$$dy = h'(x)dx \quad (10.44)$$

If the relation between X and Y is monotonic increasing, that is:

$$h(x+a) > h(x) \quad , \quad a > 0 \quad , \quad (10.45)$$

then for the probability distributions we have that:

$$P(X \leq x) = P(Y \leq h(x)) \quad . \quad (10.46)$$

Note, one must be a little careful with notation. The symbol P written here is not a function but a probability measure. The same statement in terms of the distribution functions can be written:

$$F_X(x) = G_Y(h(x)) \quad . \quad (10.47)$$

Now we have a functional relation, it is straightforward to change variable. Using the chain rule for differentiation, we can define the density for Y :

$$f_X(x) = \frac{d}{dx}F_X(x) = \frac{d}{dx}G_Y(h(x)) = \frac{dy}{dx} \times \frac{d}{dy}G_Y(y) = h'(x)g_Y(y) \quad . \quad (10.48)$$

That is the change of variable gives a change of density:

$$\boxed{g_Y(y) = \frac{f_X(x)}{h'(x)}} \quad (10.49)$$

where we need to solve the (implicit) inversion: $x = h^{-1}(y)$. In general, this is not easy to do in analytic form.

The following example shows a simple case for which this can be achieved. Consider the standard normal distribution density for the random variable, $-\infty < x < +\infty$:

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}x^2) \quad . \quad (10.50)$$

If the variable Y is related to X by $Y = e^X$, then the corresponding probability density is given by:

$$g_Y(y) = f_X(x) \times \frac{dx}{dy} = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}(\ln y)^2) \times \frac{1}{y} \quad (10.51)$$

that is the probability density for Y is given by:

$$g_Y(y) = \frac{1}{\sqrt{2\pi} y} e^{-\frac{1}{2}(\ln y)^2} \quad 0 < y < +\infty \quad . \quad (10.52)$$

This is called a *log-normal distribution*, and the shape of the distribution is shown below (10.1).

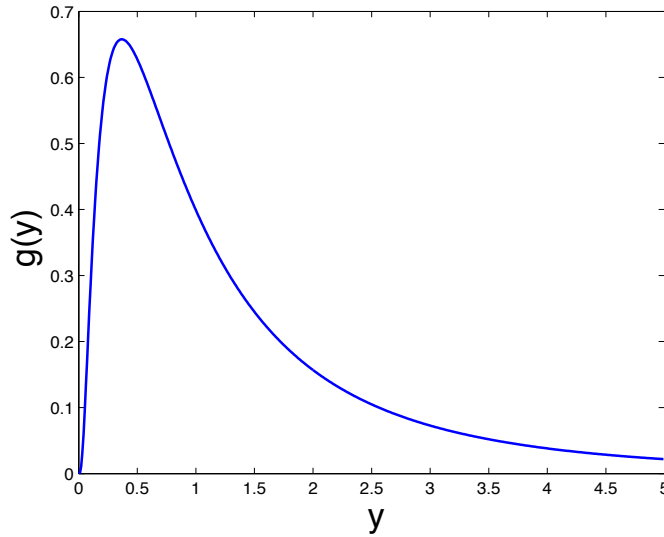


Figure 10.1: Illustration of the log-normal distribution $g(y)$ defined by equation (10.52).

10.5 Partition rule for continuous variables

Suppose we have an event A (a random variable which can be discrete or continuous). Consider a random process on the continuous variable x for which the outcome corresponds to a unique value, or range of values, of x . Then each event $X = x$ is an exclusive event, and the entire set of X -values will be a partition of the sample space. Since this is a *continuous variable* the summation over this partition is an integral. Thus the expression of the *partition rule*, in terms of a continuous variable, is as follows:

$$P(A) = \int_{-\infty}^{+\infty} P(A|X = x)f_X(x)dx \quad (10.53)$$

where the event A is *conditioned on* the values of the random variable X .

For example, suppose two cats (that we call 1 and 2) have lifetimes that are continuous (independent) variables, with exponential distributions. Let us denote the lifetime (age of cat when it dies) by $T_{1,2}$ for the two cats. Specifically, the probability that cat 1 lives longer than time, t (let's say this is measured in years), is given by the formula:

$$P(T_1 > t) = e^{-\lambda_1 t} \quad , \quad t \geq 0 \quad . \quad (10.54)$$

where $\lambda_1 > 0$ is expressed in units years^{-1} . Then we assume the probability that cat 2 lives longer than time, t , is:

$$P(T_2 > t) = e^{-\lambda_2 t} \quad , \quad t \geq 0 \quad .$$

Then the probability distributions are simply:

$$F_{T_{1,2}}(t) = P(T_{1,2} \leq t) = 1 - e^{-\lambda_{1,2}t} \quad , \quad (10.55)$$

and the corresponding densities are:

$$f_{T_{1,2}}(t) = \frac{d}{dt} [1 - e^{-\lambda_{1,2}t}] = \lambda_{1,2}e^{-\lambda_{1,2}t} \quad . \quad (10.55)$$

for $t \geq 0$, and $f_{T_{1,2}}(t) = 0$ for $t < 0$.

Recall that in both cases, the expected value of the lifetime is:

$$\mathbb{E}(T_{1,2}) = \frac{1}{\lambda_{1,2}} \quad . \quad (10.55)$$

Incidentally, these problems arise in *survival analysis* in which the parameter λ is called the *hazard rate* of the process. In the world of life insurance λ is called the force of mortality. The larger the value of λ , the shorter the expected lifetime.

The question we pose is, if the two cats are born on the same day, what is the probability that cat 1 outlives cat 2 ?

To answer this question, we use the partition rule *conditioning on* the lifetime of one of the cats: cat 2, for example. Then defining the event A as $T_1 > T_2$, that is the event that cat 1 outlives cat 2, and conditioning on the lifetime of cat 2, we have the integral:

$$P(T_1 > T_2) = \int_{-\infty}^{+\infty} P(T_1 > T_2 | T_2 = t) f_{T_2}(t) dt \quad . \quad (10.55)$$

In explicit terms this has the form:

$$P(T_1 > T_2) = \int_0^{+\infty} P(T_1 > t) \lambda_2 e^{-\lambda_2 t} dt \quad , \quad (10.55)$$

which in turn leads to:

$$P(T_1 > T_2) = \int_0^{+\infty} e^{-\lambda_1 t} \lambda_2 e^{-\lambda_2 t} dt \quad . \quad (10.55)$$

This integral is straightforward, giving the result:

$$P(T_1 > T_2) = \lambda_2 \left[\frac{e^{-\lambda_1 t - \lambda_2 t}}{-\lambda_1 - \lambda_2} \right]_0^{+\infty} = \frac{\lambda_2}{\lambda_1 + \lambda_2} \quad . \quad (10.55)$$

Is the formula correct, and does it make sense?

If $\lambda_2 \ll \lambda_1$, that is cat 2 has a much longer expected life, then it is highly unlikely that 1 will outlive 2. Comparing this deduction with the formula we obtained, we see that indeed: $P(T_1 > T_2) \rightarrow 0$. Conversely, if $\lambda_2 \gg \lambda_1$, then the short expected life of cat 2, means that cat 1 is almost certain to live longer: $P(T_1 > T_2) \rightarrow 1$. While if $\lambda_1 = \lambda_2$ we would expect that there is an equal chance of 1 outliving 2 or vice versa. So the formula makes sense in that it yields the correct result for these special cases.

10.6 Two continuous variables

Suppose we have a pair of continuous random variables, X and Y , then the joint probability density, $f_{XY}(x, y)$, is defined as:

$$P(x \leq X \leq x + dx, y \leq Y \leq y + dy) \equiv f_{XY}(x, y) dx dy \quad . \quad (10.55)$$

Then the marginal densities are defined, in direct analogy with the discrete variable definitions, as:

$$f_X(x) \equiv \int_{-\infty}^{+\infty} f_{XY}(x, y) dy \quad \text{and} \quad f_Y(y) \equiv \int_{-\infty}^{+\infty} f_{XY}(x, y) dx \quad (10.55)$$

and the probability distribution is related by:

$$P(X \leq x, Y \leq y) = F_{XY}(x, y) \equiv \int_{-\infty}^y \int_{-\infty}^x f_{XY}(x, y) dx dy \quad . \quad (10.55)$$

Following the same train of thought, the *conditional probability densities* are defined as:

$$f_{X|Y}(x|y) \equiv \frac{f_{XY}(x, y)}{f_Y(y)} \quad \text{and} \quad f_{Y|X}(y|x) \equiv \frac{f_{XY}(x, y)}{f_X(x)} \quad . \quad (10.55)$$

with the multiplication rules:

$$f_{XY}(x, y) = f_{X|Y}(x|y) f_Y(y) = f_{Y|X}(y|x) f_X(x) \quad . \quad (10.55)$$

10.6.1 Conditional expectation for continuous variables

The *conditional expectation theorem* for continuous variables takes exactly the same form as that for discrete variables, that is:

$$\boxed{\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y|X))} \quad . \quad (10.55)$$

The only point worth mentioning is that expectations involve integration rather than summation.

Proof:

The proof is analogous to that for a discrete variable; with integrations replacing the summations.

$$\mathbb{E}(Y) = \int y f_Y(y) dy = \int y \left(\int f_{XY}(x, y) dx \right) dy \quad (10.55)$$

then, using the multiplication rule (10.6) we have:

$$\int y \left(\int f_{XY}(x, y) dx \right) dy = \int y \left(\int f_{Y|X}(y|x) f_X(x) dx \right) dy \quad . \quad (10.55)$$

After some rearrangement, namely changing the order of integration (Fubini's theorem), this gives us:

$$\mathbb{E}(Y) = \int \left(\int y f_{Y|X}(y|x) dy \right) f_X(x) dx = \int (\mathbb{E}(Y|X = x)) f_X(x) dx = \mathbb{E}(\mathbb{E}(Y|X)) \quad . \quad (10.55)$$

10.7 The exponential distribution and memoryless processes

Suppose the departure time of a flight, with scheduled departure time at 10.00 a.m., has an exponential distribution. That is the true departure time T has a probability distribution:

$$P(T \leq t) = e^{-\lambda t} \quad , \quad (10.55)$$

where t is the time (in minutes) after 10.00 a.m.

The exponential distribution is unique in having *no memory*. By this we mean:

$$\boxed{P(T > t + s | T > s) = P(T > t)} \quad . \quad (10.55)$$

That is, given that the flight has not departed by a time s , the probability that it will not depart after a further time t does not depend on s . Whether any real process has such a property is open to doubt, although the process of nuclear fission, for example, is one in which this holds to almost perfection!

For the present, let's prove this *no memory* result and then discuss its meaning.

Consider an event that occurs at a random time T . By definition, an exponential distribution for this event has density:

$$f_T(t) = \lambda e^{-\lambda t} \quad . \quad (10.55)$$

Then the distribution is:

$$P(T \leq t) = \int_0^t \lambda e^{-\lambda t} dt = 1 - e^{-\lambda t} \quad . \quad (10.55)$$

So that:

$$P(T > t) = 1 - P(T \leq t) = e^{-\lambda t} \quad . \quad (10.55)$$

Thus:

$$P(T > t + s | T > s) = \frac{P(T > t + s \text{ and } T > s)}{P(T > s)} = \frac{P(T > t + s)}{P(T > s)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t} \quad . \quad (10.55)$$

10.8 Survival probabilities and waiting times

Now consider cars passing along a road. Let us suppose that the time interval between each car has an exponential distribution. Then, according to the argument above, for a pedestrian arriving at the road, at any (random or predetermined time), s , the time of the first car will have an exponential distribution, independent of the arrival time (s) of the pedestrian.

Suppose that traffic has this form, with interarrival times given by an exponential distribution with parameter, λ . We will see later that such a process is called a *Poisson process*.

A frog takes a time f to cross the road. Let us assume that if a car passes while the frog is on the road it is certain to be squashed. What is the probability that the frog survives crossing the road ?

Let's say that the frog starts at $t = 0$, we have shown above that it doesn't matter when the frog starts. Note that the first car that comes along will squash the frog, if it arrives before a time f . We are not interested in whether a second car arrives soon after, since the frog is already squashed. Therefore the probability that the frog will cross safely is the same as the probability that the first car arrives after a time f , that is,

$$P(T_1 > f) = e^{-\lambda f}$$

is the probability of survival.

We remark that as the intensity of traffic increases, $\lambda \rightarrow \infty$, then $e^{-\lambda f} \rightarrow 0$, that is the survival probability goes to zero. On the other hand as the traffic intensity goes down, $\lambda \rightarrow 0$, then $e^{-\lambda f} \rightarrow 1$, and the survival of the frog is almost certain.

Suppose the time interval between cars travelling on a long straight road has an exponential distribution, and that these times are independent. A hen wants to cross a road and it will take her a time c to do so. Furthermore, this chicken is good at judging the speed of cars, and has a long, unobstructed view view along the straight road in question. So she can estimate when it is safe for her to cross. How long does it takes the chicken to cross the road, on average ?

The passage of cars is random, and therefore the time it takes to cross is a random variable (T_c) which explains the 'on average' part of the question. So we are looking for the expected time to cross, $\mathbb{E}(T_c)$. Not surprisingly we will use the conditional expectation theorem for the calculation. Let's condition on the passage/arrival of the first car.

We have said that the arrival time is exponential The probability density for the passage of the first car is

$$f_{T_1}(t) = \lambda e^{-\lambda t} \quad , \quad t > 0 \quad . \quad (10.55)$$

There are only two cases to consider:

- $T_1 \geq c$ In this case the chicken can cross straightaway and so the time to cross is just $T_c = c$.
- $T_1 < c$
In this case, the chicken must wait for the car to pass (a time T_1) before starting to wait again. So $T_c > T_1$.

So, conditioning on the first arrival we have, using the conditional expectation theorem in the form (10.6.1):

$$\mathbb{E}(T_c) = \mathbb{E}(\mathbb{E}(T_c|T_1)) \quad (10.55)$$

that is

$$\mathbb{E}(T_c) = \int_0^\infty \mathbb{E}(T_c|T_1 = t) f_{T_1}(t) dt \quad . \quad (10.55)$$

Then breaking the integral into two sections gives:

$$\mathbb{E}(T_c) = \int_0^c \mathbb{E}(T_c|T_1 = t) f_{T_1}(t) dt + \int_c^\infty \mathbb{E}(T_c|T_1 = t) f_{T_1}(t) dt \quad . \quad (10.55)$$

The density is known, it remains to identify the conditional expectations. For the second terms $T_1 > c$, the hen crosses immediately.

$$\mathbb{E}(T_c|T_1 > c) = c \quad . \quad (10.55)$$

While for shorter arrival times, the hen cannot cross until the first car passes, and then must start waiting a random time, T_c . Then,

$$\mathbb{E}(T_c|T_1 = t) = t + \mathbb{E}(T_c) \quad , \quad t < c \quad . \quad (10.55)$$

So

$$\mathbb{E}(T_c) = \int_0^c (t + \mathbb{E}(T_c)) \lambda e^{-\lambda t} dt + \int_c^\infty c \lambda e^{-\lambda t} dt \quad . \quad (10.55)$$

The integrals can be done by standard means:

$$\int_0^c (t + \mathbb{E}(T_c)) \lambda e^{-\lambda t} dt = -ce^{-\lambda c} + \frac{1}{\lambda} [1 - e^{-\lambda c}] + \mathbb{E}(T_c) [1 - e^{-\lambda c}] \quad . \quad (10.55)$$

$$\int_c^\infty c \lambda e^{-\lambda t} dt = ce^{-\lambda c} \quad . \quad (10.55)$$

Then combining these expressions gives:

$$\mathbb{E}(T_c) = -ce^{-\lambda c} + \frac{1}{\lambda} [1 - e^{-\lambda c}] + \mathbb{E}(T_c) [1 - e^{-\lambda c}] + ce^{-\lambda c} \quad . \quad (10.55)$$

After some simple rearrangement we get:

$$\mathbb{E}(T_c) = \frac{1}{\lambda} [e^{\lambda c} - 1] \quad . \quad (10.55)$$

Again, we should check that the answer makes sense. Suppose that the traffic is extremely intense, $\lambda \rightarrow \infty$. Then the expected time to cross will be infinite.

The other extreme, very infrequent traffic, $\lambda \rightarrow 0$, we find (using L'Hôpital's rule) the limit:

$$\lim_{\lambda \rightarrow 0} \mathbb{E}(T_c) = c \quad (10.55)$$

as we would expect. Since in this case, the first car takes an infinite amount of time to arrive, and the hen can walk straight across the road without having to wait. So the time it takes to cross is simply, c .