Lecture 4: Combining and Manipulating Distributions

IB Paper 7: Probability and Statistics

Carl Edward Rasmussen

Department of Engineering, University of Cambridge

February 10th, 2009

Waiting times

The bus arrives on average every 15 minutes. Compare the average waiting time for people arriving randomly if buses 1) arrive regularly, 2) arrive randomly.

1) Buses arrive regularly every 15 minutes. The waiting time is uniformly distributed in [0; 15], $p(x) = 1/15, 0 \leqslant x \leqslant 15$ and zero otherwise. The average is

$$\mathbb{E}[X] \;=\; \frac{1}{15}\int_0^{15} x dx \;=\; \frac{1}{15}\big[\tfrac{1}{2}x^2\big]_0^{15} \;=\; \frac{15}{2} \text{ minutes.}$$

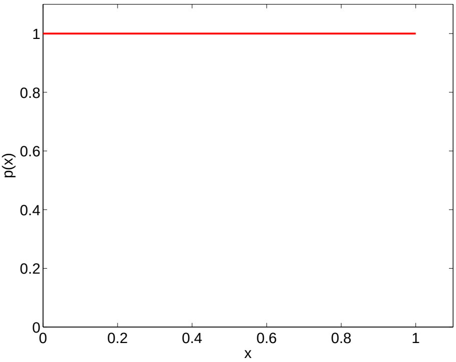
2) Buses arrive randomly, according to a Poisson, $\text{Po}(\lambda = 1/15)$. Therefore the waiting time to the next event is exponentially distributed $\text{Ex}(\lambda = 1/15)$ with expectation

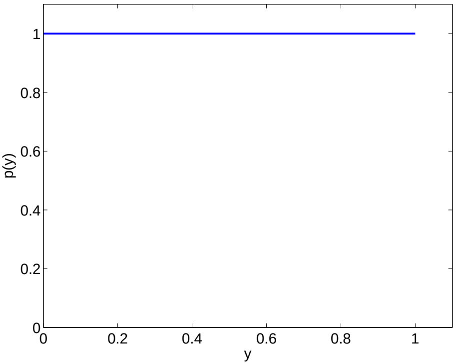
$$\begin{aligned}\mathbb{E}[X] \;&=\; \int_0^\infty x\lambda \exp(-\lambda x) dx \;=\; -\big[x\exp(-\lambda x)\big]_0^\infty + \int_0^\infty \exp(-\lambda x) dx \\ &=\; -\big[\frac{1}{\lambda}\exp(-\lambda x)\big]_0^\infty \;=\; \frac{1}{\lambda} \;=\; 15 \text{ minutes}\end{aligned}$$

How is it possible, that when the average bus arrival rate is the same, the average waiting time differs by a factor of 2?

Combining Probabilities

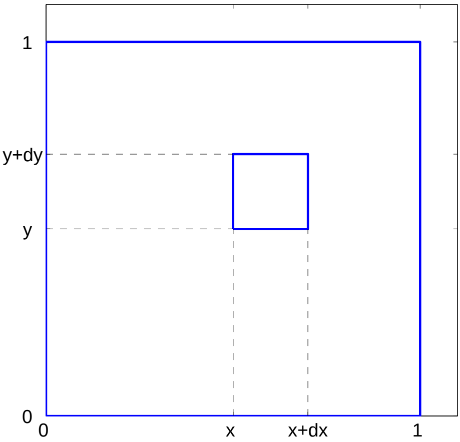
Assume the variables X and Y independent, uniformly distributed





We wish to compute the distribution of the sum of the two random variables, $S = X + Y$.

The joint probability is uniform on the unit square

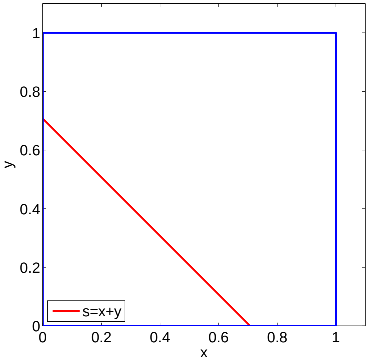


The probability of $x \leqslant X \leqslant x + dx$ and $y \leqslant Y \leqslant y + dy$.

The joint probability is $p(x, y) = 1$. For independent variables, $p(x, y) = p(x)p(y)$.

Distribution of the sum based on $F(s)$

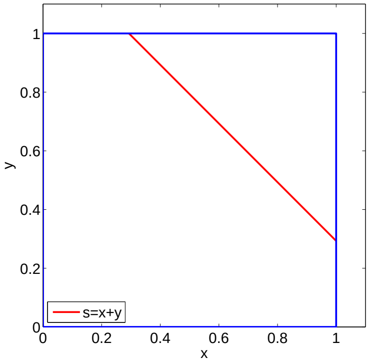
Assume first $0 \leqslant s \leqslant 1$. We have $F(s) = p(S < s) = s^2/2$.



Therefore

$$p(s) = \frac{dF(s)}{ds} = s, \text{ where } 0 \leqslant s \leqslant 1.$$

For $1 \leqslant s \leqslant 2$, we have $F(s) = 1 - (2-s)^2/2$.



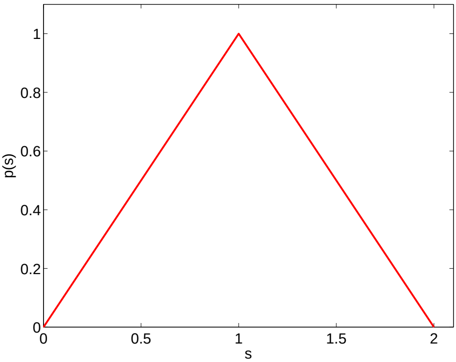
Therefore

$$p(s) \;=\; \frac{dF(s)}{ds} \;=\; 2-s, \;\; \text{where} \;\; 1 \leqslant s \leqslant 2.$$

The distribution of the sum

Thus

$$p(s) = \begin{cases} s & 0 \leqslant s \leqslant 1 \\ 2 - s & 1 \leqslant s \leqslant 2 \end{cases}$$



It is more likely that the sum is around 1, than that it is close to 0 or 2.

Means and Variances of sums of independent variables

The mean is

$$\mathbb{E}[S] \;=\; \int_0^2 sp(s)ds \;=\; \int_0^1 s^2 ds + \int_1^2 s(2-s)ds \;=\; [\frac{s^3}{3}]_0^1 + [s^2 - \frac{s^3}{3}]_1^2 \;=\; 1.$$

And second moment

$$\mathbb{E}[S^2] \;=\; \int_0^2 s^2 p(s)ds \;=\; \int_0^1 s^3 ds + \int_1^2 s^2(2-s)ds \;=\; [\frac{s^4}{4}]_0^1 + [\frac{2s^3}{3} - \frac{s^4}{4}]_1^2 \;=\; \frac{7}{6},$$

so the variance is $\mathbb{E}[S^2] - \mathbb{E}^2[S] = 1/6$.

Notice:

- The mean of the sum is the sum of the means
- The variance of the sum is the sum of the variances (holds for sums of *independent* variables).

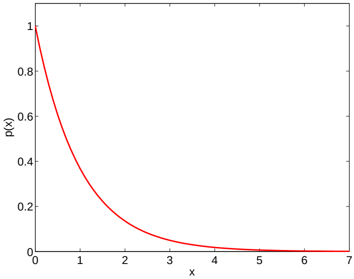
Sums of Variables based on the densities

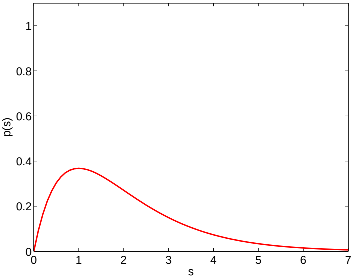
Assume that X and Y are independent, and $S = X + Y$. What is the distribution of S?

$$p(S = s) \;=\; \int p(X = x)p(Y = s - x)dx$$

Example: Assume $X = \text{Ex}(1)$, i.e. $p(x) = \exp(-x)$, and $Y = \text{Ex}(1)$. Then, $S = X + Y$ has distribution

$$\begin{aligned} p(S = s) \;&=\; \int_0^\infty p(X = x)p(Y = s - x)dx \;=\; \int_0^s \exp(-x)\exp(-s + x)dx \\ &=\; \exp(-s)[x]_0^s \;=\; s\exp(-s). \end{aligned}$$





More on Sums

Example: $p(x)$ and $p(y)$ are independent Gaussian $\mathrm{N}(\mu_x, \sigma_x^2)$ and $\mathrm{N}(\mu_y, \sigma_y^2)$. The sum $S = X + Y$ is also Gaussian $\mathrm{N}(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$.

Example: The difference $T = X - Y$ is also Gaussian $\mathrm{N}(\mu_x - \mu_y, \sigma_x^2 + \sigma_y^2)$.

Notice:

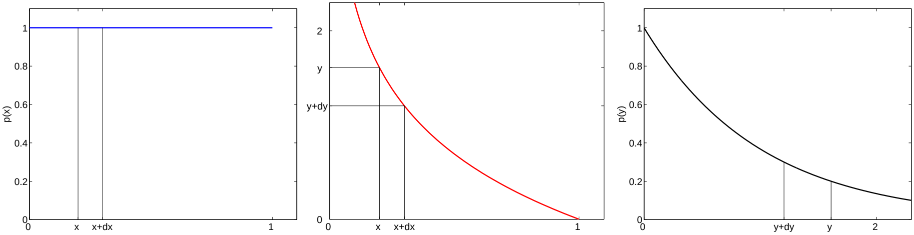
- The sum of two independent Gaussians is again Gaussian
- But, as we saw before, the sum of Exponentials is not Exponential

Gaussians are said to be *closed under addition*.

Because of the central limit theorem, distributions become *more Gaussian* when you add them.

Transformations of continuous variables

The distribution of x is $p(x)$. What is the distribution of $y = f(x)$?



We want the probability of an event in the old variables x to be equal to the probability in the new variables

$$p(x)dx \;=\; p(y)dy.$$

The Jacobian

Formally, we have

$$p(y) \;=\; p(x)\left|\frac{dx}{dy}\right|,$$

where $|\frac{dx}{dy}|$ is called the Jacobian.

Example: $p(x)$ is Gaussian $\mathrm{N}(\mu, \sigma^2)$, i.e.

$$p(x) \;=\; \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{1}{2\sigma^2}(x-\mu)^2).$$

What is the distribution of $p(y)$ of $y = \alpha x + \beta$, a linear function of x?

$$x \;=\; \frac{y-\beta}{\alpha}, \quad \text{so} \quad |\frac{dx}{dy}| = \frac{1}{\alpha}, \quad \text{and}$$
$$p(y) = \frac{1}{\alpha}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\Big(-\frac{1}{2\sigma^2}\big(\frac{y-\beta}{\alpha}-\mu\big)^2\Big),$$

which we recognize exactly as $p(y)$ is $\mathrm{N}(\beta + \alpha\mu, \alpha^2\sigma^2)$.

Linear transformations

Linear transformations don't change the shape of the distribution.

- Adding a constant simply shifts the mean of the distribution by that constant.
- Multiplication by a constant
  - scales the mean by the constant, and
  - scales the variance by the *square* of that constant.

The Jacobian for a non-linear transformation

For a linear transformation the Jacobian is just a constant, which makes sure the the probabilities in the new variables normalize.

For non-linear transformations, the Jacobian is more complicated.

Example: $p(x)$ is uniform $[0;1]$. What is the distribution of $y = -\log(x)$?

Since $x = \exp(-y)$, the Jacobian is $|-\exp(-y)| = \exp(-y)$, so

$$p(y) \;=\; p(x)\exp(-y) \;=\; \exp(-y).$$

So, y follows an exponential distribution, $\mathrm{Ex}(1)$.