Lecture 23: Chebychev theorem

In this lecture we look at more probability distributions and prove the fantastically useful Chebychev's theorem.

Remember that a **continuous probability density** is a nonnegative function f such that $\int_R f(x)\,dx = 1$. A random variable X has this probability density if

$$\mathrm{P}[X \in [a,b]\,] = \int_a^b f(x)\,dx$$

for all intervals $[a,b]$.

If we know the probability density of a random variable, we can compute all the important quantities like the expectation or the variance.

If X has the probability density f, then $m = \mathrm{E}[X] = \int x f(x)\,dx$ and $\mathrm{Var}[X] = \int (x-m)^2 f(x)\,dx$.

The **distribution function** of a random variable with probability density f is defined as

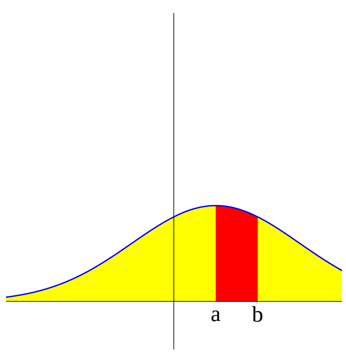
$$F(s) = \int_{-\infty}^{x} f(x)\,dx = \mathrm{P}[X \leq s]\,.$$

By definition F is a monotone function: $F(b) \geq F(a)$ for $b \geq a$. One abbreviates the probability density function with PDF and the distribution function with CDF which abbreviates cumulative distribution function.

1 The most important distribution on the real line is the **normal distribution**

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}\,.$$

It has mean m and standard deviation σ. This is a probability measure because after a change of variables $y = (x-m)/(\sqrt{2\sigma})$, the integral $\int_{-\infty}^{\infty} f(x)\,dx$ becomes $\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-y^2}\,dy = 1$.



2 The most important distribution on the positive real line is the **exponential distribution**

$$f(x) = \lambda e^{-\lambda x}\,.$$

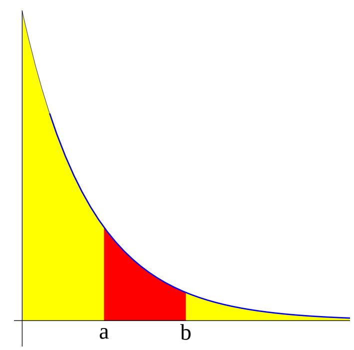
Lets compute its mean:

$$m = \int_0^\infty x f(x)\,dx = \frac{1}{\lambda}\,.$$

From $\lambda \int_0^\infty x^2 \exp(-\lambda x)\,dx = 2/\lambda^2$, we get the variance

$$2/\lambda^2 - 1/\lambda^2 = 1/\lambda^2$$

and the standard deviation $1/\lambda$.



3 The most important distribution on a finite interval $[a,b]$ is the **uniform distribution**

$$f(x) = 1_{[a,b]}\frac{1}{b-a}\,,$$

where 1_I is the characteristic function

$$1_I(x) = \begin{cases} 1 & \mathrm{x} \in \mathrm{I} \\ 0 & \mathrm{x} \notin \mathrm{I} \end{cases}\,.$$

The following theorem is very important for estimation purposes. Despite the simplicity of its proof it has a lot of applications:

Chebychev theorem If X is a random variable with finite variance, then

$$\mathrm{P}[|X - \mathrm{E}[X]| \geq c] \leq \frac{\mathrm{Var}[X]}{c^2}\,.$$

Proof. The random variable $Y = X - \mathrm{E}[X]$ has zero mean and the same variance. We need only to show $\mathrm{P}[|Y| \geq c] \leq \frac{\mathrm{Var}[Y]}{c^2}$. Taking the expectation of the inequality

$$c^2 1_{\{|Y| \geq c\}} \leq Y^2$$

gives

$$c^2 \mathrm{P}[|Y| \geq c] \leq \mathrm{E}[Y^2] = \mathrm{Var}[Y]$$

finishing the proof.

The theorem also gives more meaning to the notion "Variance" as a measure for the deviation from the mean. The following example is similar to the one section 11.6 of Cliff's notes:

4 A die is rolled 144 times. What is the probability to see 50 or more times the number 6 shows up? Let X be the random variable which counts the number of times, the number 6 appears. This random variable has a binomial distribution with $p = 1/6$ and $n = 144$. It has the expectation $\mathrm{E}[X] = np = 144/6 = 24$ and the variance $\mathrm{Var}[X] = np(1-p) = 20$. Setting $c = (50-24) = 26$ in Chebychev, we get $\mathrm{P}[|X - 24| \geq 26] \leq 20/26^2 \sim 0.0296....$ The chance is smaller than 3 percent. The actual value $\sum_{k=50}^{144} \binom{144}{k} p^k (1-p)^{144-k} \sim 1.17 \cdot 10^{-7}$ is much smaller. Chebychev does not necessarily give good estimates, but it is a handy and univeral "rule of thumb".

Finally, lets look at a practical application of the use of the cumulative distribution function. It is the task to **generate random variables with a given distribution**:

5 Assume we want to generate random variables X with a given distribution function F. Then $Y = F(X)$ has the uniform distribution on $[0,1]$. We can reverse this. If we want to produce random variables with a distribution function F, just take a random variable Y with uniform distribution on $[0,1]$ and define $X = F^{-1}(Y)$. This random variable has the distribution function F because $\{X \in [a,b]\ \} = \{F^{-1}(Y) \in [a,b]\ \} = \{Y \in F([a,b])\ \} = \{Y \in [F(a), F(b)]\} = F(b) - F(a)$. We see that we need only to have a random number generator which produces uniformly distributed random variables in $[0,1]$ to get a random number generator for a given continuous distribution. A computer scientist implementing random processes on the computer only needs to have access to a random number generator producing uniformly distributed random numbers. The later are provided in **any** programming language which deserves this name.

To generate random variables with cumulative distribution function F, we produce random variables X with uniform distribution in $[0,1]$ and form $Y = F^{-1}(X)$.

With computer algebra systems

1) In Mathematica, you can generate random variables with a certain distribution with a command like in the following example:

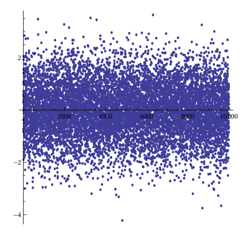
```
X:=Random[NormalDistribution[0,1]]
ListPlot[Table[X,{10000}],PlotRange->All]
```

2) Here is how to access the **probability density function** (PDF)

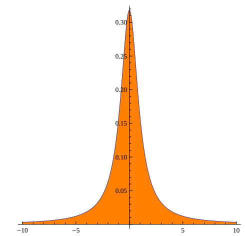
```
f=PDF[CauchyDistribution[0,1]];
S=Plot[f[x],{x,-10,10},PlotRange->All,Filling->Axis]
```

3) And the **cumulative probability distribution** (CDF)

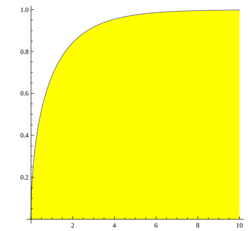
```
f=CDF[ChiSquareDistribution[1]];
S=Plot[f[x],{x,0,10},PlotRange->All,Filling->Bottom]
```



Random numbers



The PDF



The CDF

Homework due March 30, 2011

1 The random variable X has a normal distribution with standard deviation 2 and mean 5. Estimate the probability that $|X - 5| > 3$.

2 Estimate the probability of the event $X > 10$ for a Poisson distributed random variable X with mean 4.

3 a) Verify that $\phi(x) = \tan(x\pi)$ maps the interval $[0,1]$ onto the real line so that its inverse $F(y) = \arctan(y)/\pi$ is a map from R to $[0,1]$.
b) Show that $f = F'(y) = \frac{1}{\pi}\frac{1}{1+y^2}$.
c) Assume we have random numbers in $[0,1]$ handy and want to random variables which have the probability density f. How do we achieve this?
d) The mean $\int_{-\infty}^{\infty} x f(x)\, dx$ does not exist as an indefinite integral but can be assigned the value 0 by taking the limit $\int_{-R}^{R} x f(x)\, dx = 0$ for $R \to \infty$. Is it possible to assign a value to the variance $\int_{-\infty}^{\infty} x^2 f(x)\, dx$?.

The probability distribution with density

$$\frac{1}{\pi}\frac{1}{1+y^2}$$

which appeared in this homework problem is called the **Cauchy distribution**. Physicists call it the **Cauchy-Lorentz distribution**.

Why is the Cauchy distribution natural? As one can deduce from the homework, if you chose a random point P on the unit circle, then the slope of the line OP has a Cauchy distribution. Instead of the circle, we can take a rotationally symmetric probability distribution like the Gaussian with probability measure $P[A] = \int\int_A e^{-x^2-y^2}/\pi\ dxdy$ on the plane. Random points can be written as (X,Y) where both X,Y have the normal distribution with density $e^{-x^2}/\sqrt{\pi}$. We have just shown

If we take independent Gaussian random variables X,Y of zero mean and with the same variance and form $Z = X/Y$, then the random variable Z has the Cauchy distribution.

Now, it becomes clear why the distribution appears so often. Comparing quantities is often done by looking at their ratio X/Y. Since the normal distribution is so prevalent, there is no surprise, that the Cauchy distribution also appears so often in applications.