

# Stochastic models

Forrest W. Crawford



# Stochastic models

Stochastic means random.

## Differences from deterministic modeling

# Random variables

## Probability

Probability  $\Pr(\cdot)$  is a function whose argument is the value a random variable can take on, which returns a positive number less than or equal to 1. The set of all such values describes the probability distribution.

We usually describe random variables in terms of their probability distribution. For example, consider the random variable  $X$  defined by  $\Pr(X = 0) = 1/2$  and  $\Pr(X = 1) = 1/2$ . What might  $X$  refer to?

## Cumulative probability distribution function

All random variables on the real line have a cumulative distribution function (CDF)

$$F(x) = \Pr(X < x)$$

## Density

$f(x)$  is derivative of CDF, where it exists

$$f(x) = \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h}$$



## Expectation

The expected value of a random variable is its “average” value.

$$E[X] = \sum_x x \Pr(X = x)$$

or

$$E[X] = \int_{-\infty}^{\infty} x dP(x)$$

average of values  $x$  that  $X$  can take on, weighted by their probability.

Can a random variable have expectation equal to a value that  $X$  can never take? Yes. In the example above  $\Pr(X = x) = 1/2$  for  $x = 1$  and  $x = 0$ . But  $E[X] = 1/2$ .

## Independence

Two random variables are independent if their joint probability factorizes into the product of their marginal probabilities,

$$\Pr(X = x, Y = y) = \Pr(X = x) \Pr(Y = y)$$

e.g. independent coin flips, die rolls, etc.

# Common probability distributions

## Bernoulli

This is the classic coin-flipping distribution. The random variable  $X$  has Bernoulli( $p$ ) distribution if  $\Pr(X = 1) = p$  and  $\Pr(X = 0) = 1 - p$ .

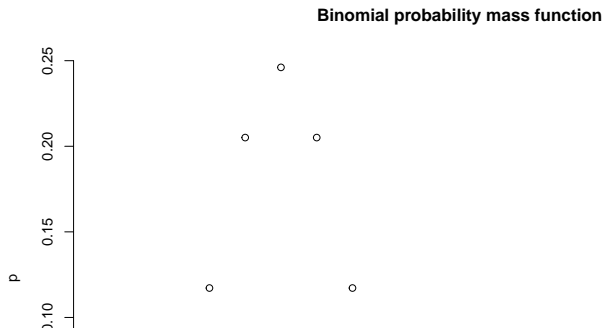
# Binomial

sum of bernoullis For Binomial( $n, p$ ),

$$\Pr(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

where  $0 \leq k \leq n$ .

```
ks = 0:20  
p = dbinom(ks,size=10,p=1/2)  
plot(ks,p, bty="n", main="Binomial probability mass function")
```



# Geometric

number of bernoulli trials to get a particular number of successes

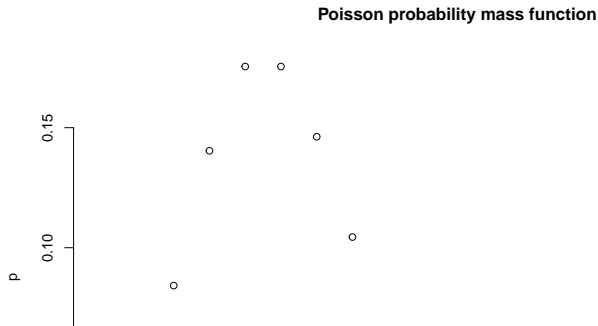
# Negative binomial

## Poisson

A common model for count data is

$$\Pr(X = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$$

```
ks = 0:20  
lam = 5  
p = dpois(ks,lam)  
plot(ks,p, bty="n", main="Poisson probability mass function")
```





# Continuous

## Uniform

Usually defined on  $[0, 1]$ .  $X \sim \text{Unif}(0, 1)$  means the density is

$$f(x) = 1$$

## Exponential

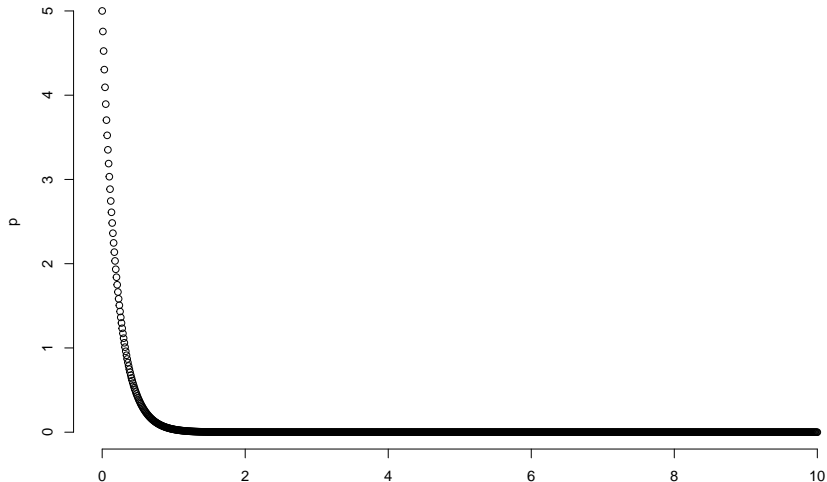
$X$  has exponential distribution if

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

$$F(x) = 1 - e^{-\lambda t}$$

```
xs = seq(0,10,by=0.01)
lam = 5
p = dexp(xs, rate=lam)
plot(xs, p, bty="n", main="Exponential probability density")
```

Exponential probability density



Memoryless property:

$$\Pr(X > t+h|X > t) = \frac{\Pr(X > t+h)}{\Pr(X > t)} = \frac{e^{-\lambda(t+h)}}{e^{-\lambda t}} = e^{-\lambda h} = \Pr(X > h)$$

Normal

## Others

Gamma, lognormal, etc. Survival distributions There are lots more.

# Structural probability models

So-called statistical models are stochastic. Consider the Normal/Gaussian linear model

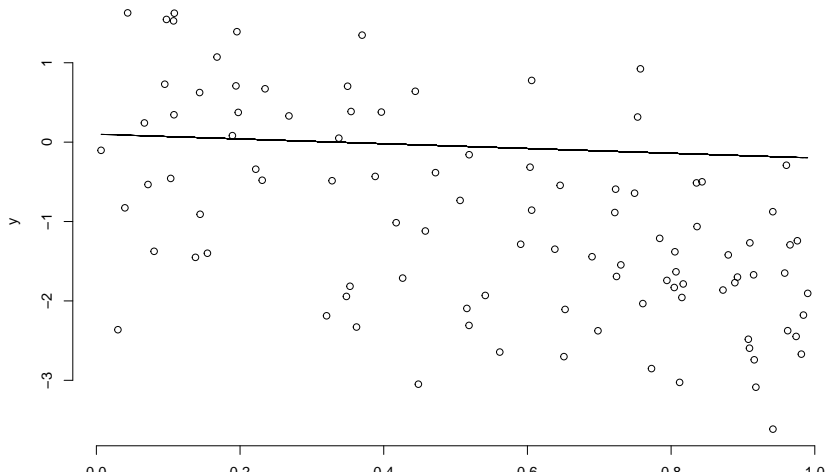
$$y = \alpha + \beta x + \epsilon$$

where  $\epsilon \sim N(0, \sigma^2)$ .

This model is stochastic because  $\epsilon$  is random. It could be that  $x$  is assigned randomly.



```
n = 100  
eps = rnorm(100)  
x = runif(n)  
y = 0.1 - 2*x + eps  
plot(x,y, bty="n")  
lines(x,0.1-0.3*x)
```



# Stochastic mechanistic models

We described deterministic systems in terms of their rates earlier. But a stochastic model produces different output each time we run it. How can we describe its rates of change when it's different every time?

Answer: describe the rate of change of its probabilities or moments. To do this, we need to set up ODEs:

$$P_{ij}(t) = \Pr(X(t+h) = j | X(t) = i)$$

## Kolmogorov equations. . . [fix]

$$\frac{dP_{ij}(t)}{dt} = \lambda_{j-1}P_{i,j-1}(t) - \lambda_j P_{ij}(t)$$

Master equation for rxn systems

Relationship to “statistical models”

Advanced: moment generating functions.

# modeling stochastic processes with rates

Modeling a waiting time with the exponential distribution

Poisson process

Time-varying poisson process

Counting processes

## Pure birth process

Let  $X(t)$  be the number of individuals in a population, and let  $X(0) = x_0 > 0$  be the initial number. Suppose the population gains another individual with the rate

$$\lambda_k = k\lambda$$

The expected number of individuals at time  $t$  is

$$E[X(t)] = x_0 e^{\lambda t}$$

Exponential growth!

## Pure death process

Suppose the population has  $X(0)$  individuals initially, and loses one with rate

$$\mu_k = k\mu$$

The expected number of individuals at time  $t$  is

$$E[X(t)] = x_0 e^{-\mu t}$$

Exponential decay!

## Birth-death

$$\lambda_k = k\lambda$$

$$\mu_k = k_m u$$

The expected number of individuals at time  $t$  is

$$E[X(t)] = x_0 e^{(\lambda - \mu)t}$$

## Birth-death immigration/emigration

$$\lambda_k = \alpha + k\lambda$$

$$\mu_k = \delta + k_m u$$

Expectations are complicated!



## Hawkes self-exciting Poisson process

$$\lambda(t) = \alpha(t) + \sum_{i=1}^{N(t)} M_i e^{-\gamma(t-t_i)}$$

# Simulating stochastic population models

## Gillespie algorithm

This is for forward simulation of a compartmental model. Generally will not work for individualistic (agen-based) models.

Collapsing model compartments for easier simulation

Representing heterogeneity

## Agent-based (micro-)simulation

Exercise: single household/network SI epidemic model agent-based sim with covariates

```
n = 10
p = 3

a = 0.2
b = c(-1,0.3,0.1)
g = c(-0.2,1,0.3)

x = matrix(rnorm(n*p,0,1), nrow=n)
tmax = 1

infectives = c()
susceptibles = 1:n

shaz = a + exp(z %*% b) # this is a vector of length n
ihaz = exp(z %*% g) # this is a vector of length n
```

[\[Back: Deterministic Models\]](#) [\[Home\]](#) [\[Next: Agent-based models\]](#)

## References