1. Intercepts in regression.
2. LTS.
3. Inhomogeneous Poisson processes.
4. Non-Poisson point processes.
5. Poisson cluster processes, self-exciting point processes, Neyman-Scott processes, Thomas processes, and Matern processes.

1. Intercept.
Generally, it is best to keep the intercept in a regression model, whether it is statistically significantly different from zero or not.

2. LTS with more than 1 explanatory variable.

```
x1 = 10*rnorm(1000)
x2 = 10*rnorm(1000)
y1 = 1+2*x1+3*x2 + 4*rnorm(1000)
plot(x1,y1)
a = lm(y1 ~ x1+x2)
summary(a)
abline(a)
x1[1:3] = x1[1:3]+200
x2[1:3] = x2[1:3]+200
plot(x1,y1)
abline(1,2,lty=1)
a = lm(y1 ~ x1+x2)
a$coef
abline(a,lty=2)
library(MASS)
a2 = ltsreg(y1~ x1+x2,.95)
a2$coef
```

```
abline(a2,lty=3)
legend(100,1,legend = c("truth","OLS","LTS"),
lty=c(1,2,3))
plot(x2,y1)
abline(1,3,lty=1)
abline(a$coef[1],a$coef[3],lty=2)
abline(a2$coef[1],a2$coef[3],lty=3)
legend(100,1,legend = c("truth","OLS","LTS"),
lty=c(1,2,3))
```

3. What's an example of an *inhomogeneous* Poisson process?

Don't worry about point process topics, like intensity estimation, for your projects! You are using regression data for your projects, not point process data, so those topics do not apply.

To specify a Poisson process, all you need to do is specify the rate $\lambda(t)$. For instance, suppose the space S is a portion of the real time line, [0,10]. You could have $\lambda(t) = t+2.5$, or $t^2 + 4t + 2$. In both of these cases we'd expect few points near the origin, and lots more points as time increases.

What about $t^2 + 4t - 2$?

The rate can't be negative.

If instead $\lambda(t) = 2.5$, then it's a homogeneous Poisson process, and we expect 2.5 points per unit time.

```
## simulate a homogeneous Poisson process with rate 5
## from time 0 to time 10.
n = rpois(1,50)
t = runif(n)*10
```

```
plot(c(0,10),c(0,10),type="n",xlab="time",ylab="intensity")
points(t,rep(1,n))
lines(c(0,10),c(5,5))
## simulate an inhomogeneous Poisson process with rate
## 1+t between time 0 and 10
## the integral from 0 to 10 of 1+t = 10 + 50 = 60.
## the maximum rate is 11 between time 0 and time 10.
n = rpois(1,110)
t1 = runif(n)*10 ## a Poisson process with rate 11.
t2 = t1[runif(n)< (1+t1)/11]
## keep points with prob. lambda/11
n2 = length(t2)
plot(c(0,10),c(0,11),type="n",xlab="time",ylab="intensity")
points(t2,rep(1,n2))
lines(c(0,10),c(1,11))
```

4) What's an example of a non-Poisson point process?

It may be self-exciting, or self-correcting. That is, the occurrance of points in some time interval B1 may influence the number of points that will occur in a nearby interval B2. Remember, in order for the point process to be Poisson, N(B1) must be independent of N(B2), for disjoint sets B1 and B2.
Self-exciting = clustered.
Self-correcting = inhibitory.

What's an example of a self-exciting point process?

Poisson cluster process: start with a Poisson process,

and for each point in it, consider that point a ``parent''. For each parent, generate a random number of ``children'', so that the parents' numbers of children are iid from some probability distribution P. Let the spatial locations of the children be dispersed around the parent, independently of each other, according to some spatial density f.
Then look at the final map of all the children.
This is a self-exciting process: if you know there are points in B1, then it's more likely that there will be points in a nearby region B2.

A process could also be self-correcting, or inhibitory.

5. Neyman-Scott processes, Thomas processes, and Matern inhibition processes.
For a standard Neyman-Scott process, a fixed number of offspring points are uniformly spread in a ball around each parent.
For a Thomas process, the number of offspring points is random and the offspring points are normally distributed around their parents.
With a Matern inhibition process, you start with a homogeneous Poisson process, and pick some number δ. Find all pairs of points that are within δ of one another, and delete all such points. The remaining points must be well-dispersed. If you know there's a point in B1, then that tells you it is less likely that there's a point in a nearby region B2.

```
library(spatstat)
win1 = owin(c(0,10),c(0,10))
```

```
## Neyman-Scott process
nclust = function(x0, y0, radius, n) {
   return(runifdisc(n, radius, centre=c(x0, y0)))
          }
x = rNeymanScott(.1, 0.2, nclust, radius=0.2, n=7, win1)
plot(x,pch=".",main="Neyman Scott process")

## Thomas process
x = rThomas(.1,.1,7,win1)
## intensity, SD of distances, exp. points per cluster,
plot(x,pch=".",main="Thomas process")

## Matern inhibition process
x = rMaternI(1,.5,win1)
## intensity of the original Poisson process,
## range of exclusion)
plot(x,main="Matern process")
x2 = rMaternI(1,0,win1)
plot(x2,main="Poisson process")
```