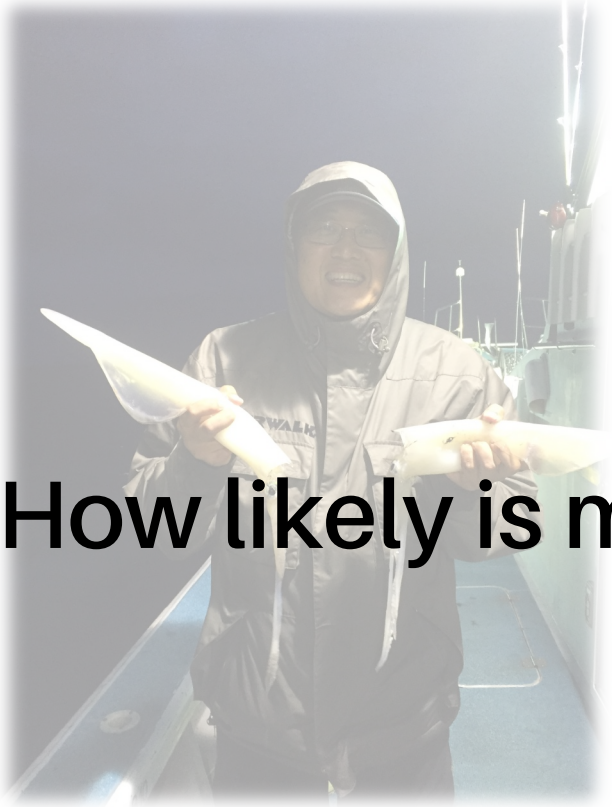
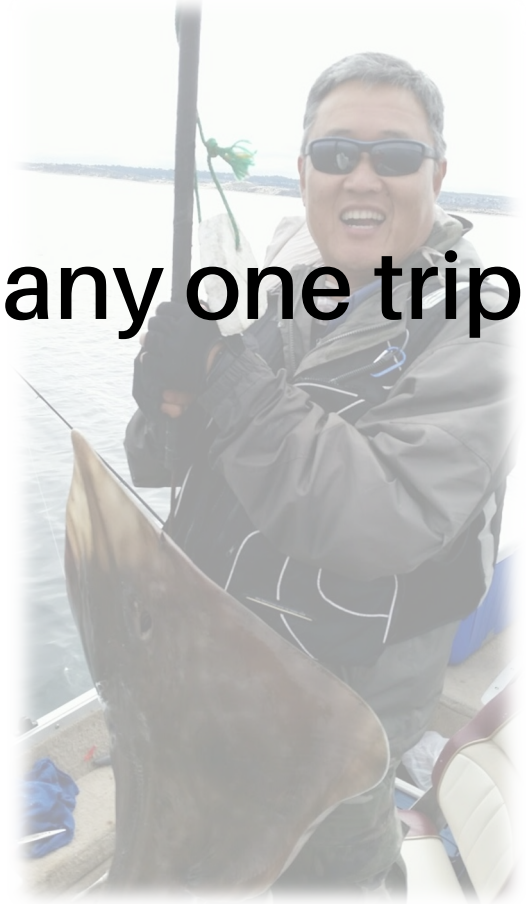
Poisson Regression

PSY504 Spring 2023


Sarah Joo

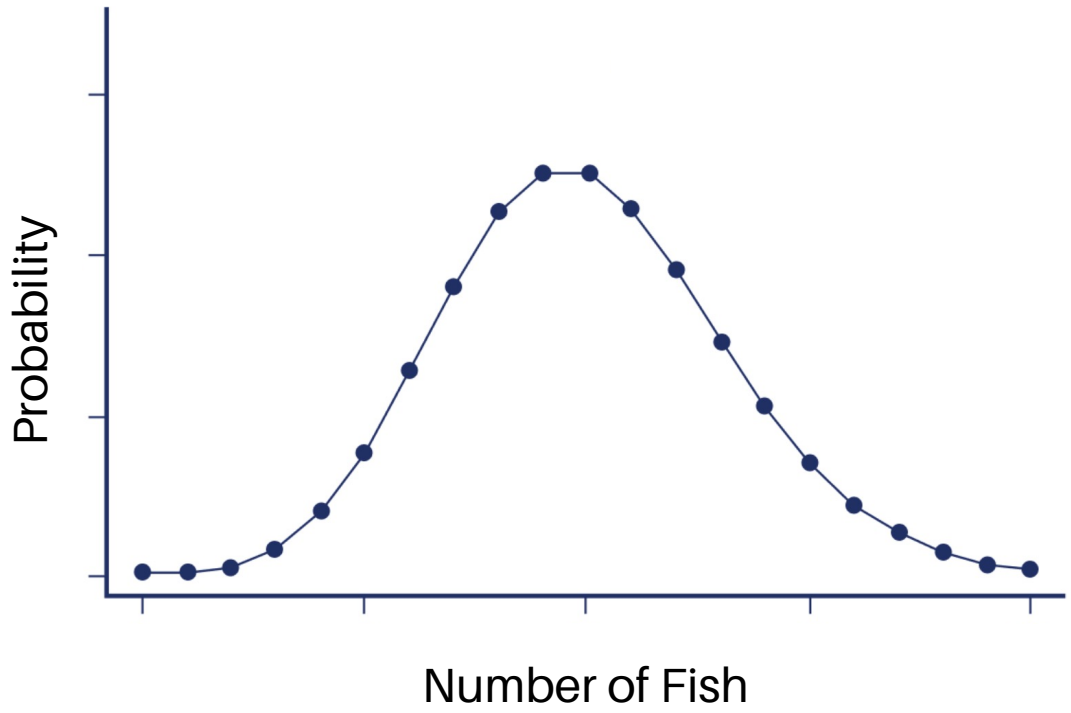




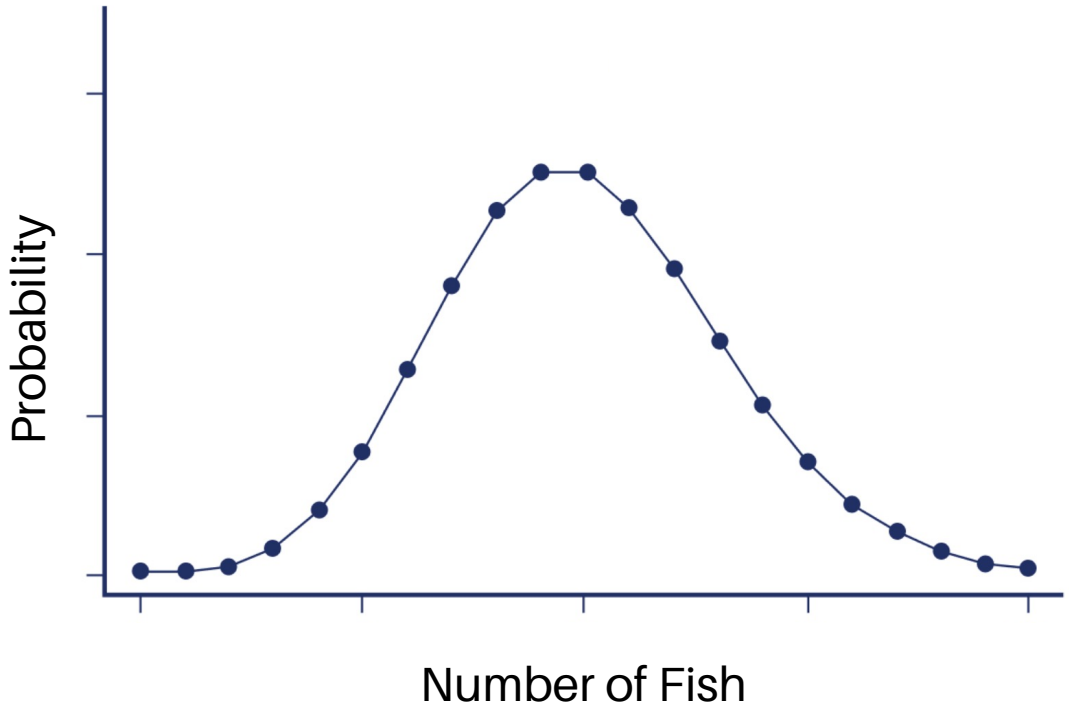
How likely is my dad to catch [#] of fish on any one trip?

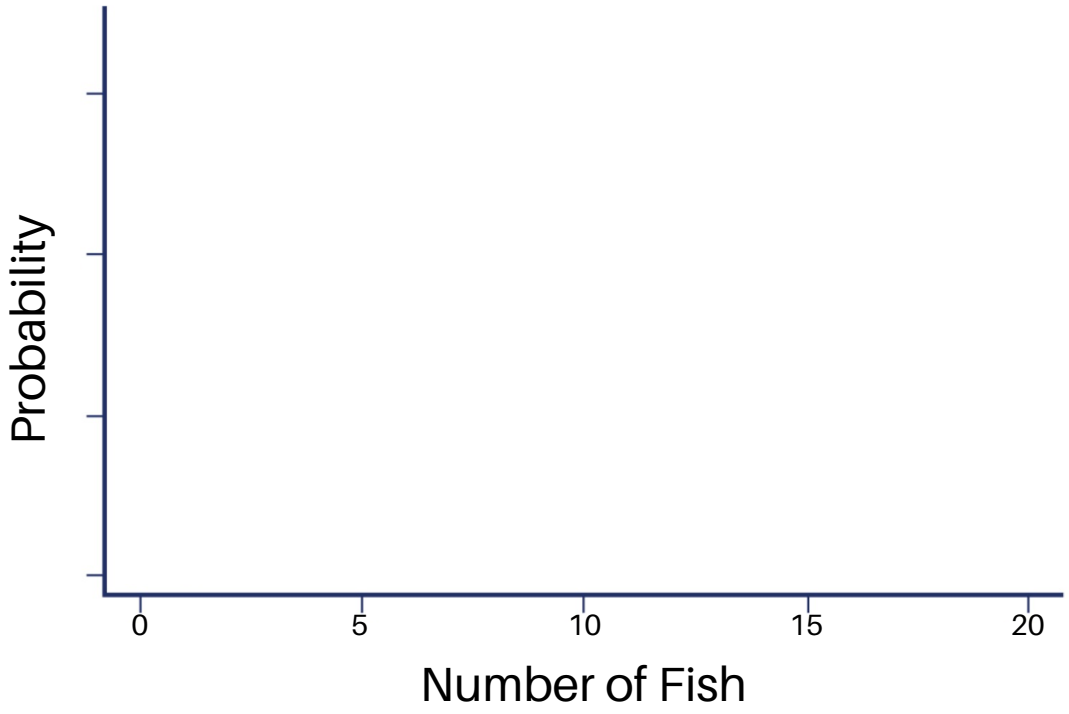
How likely is my dad to catch [#] of fish on any one trip?

How likely is my dad to catch [#] of fish on any one trip?

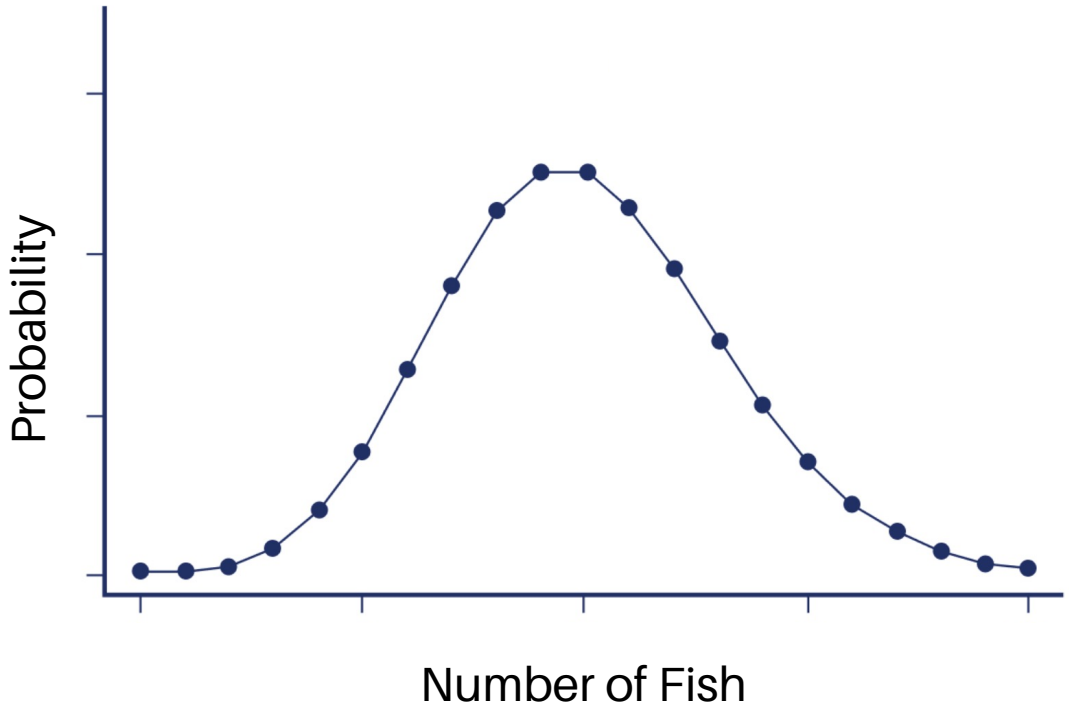


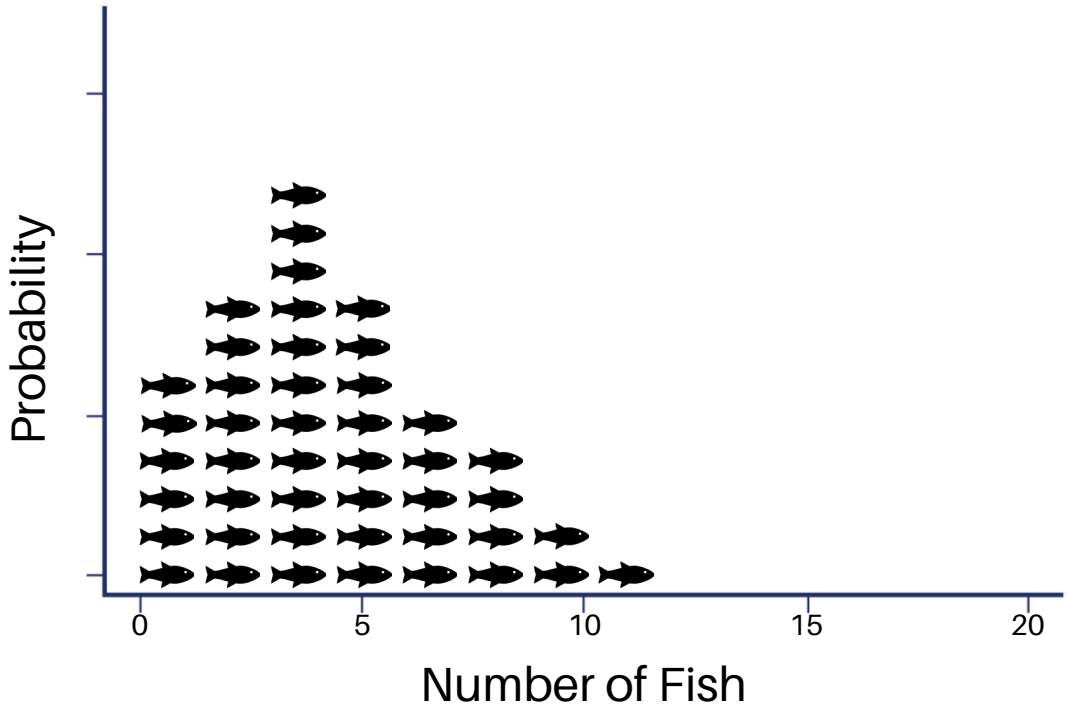
How likely is my dad to catch [#] of fish on any one trip?





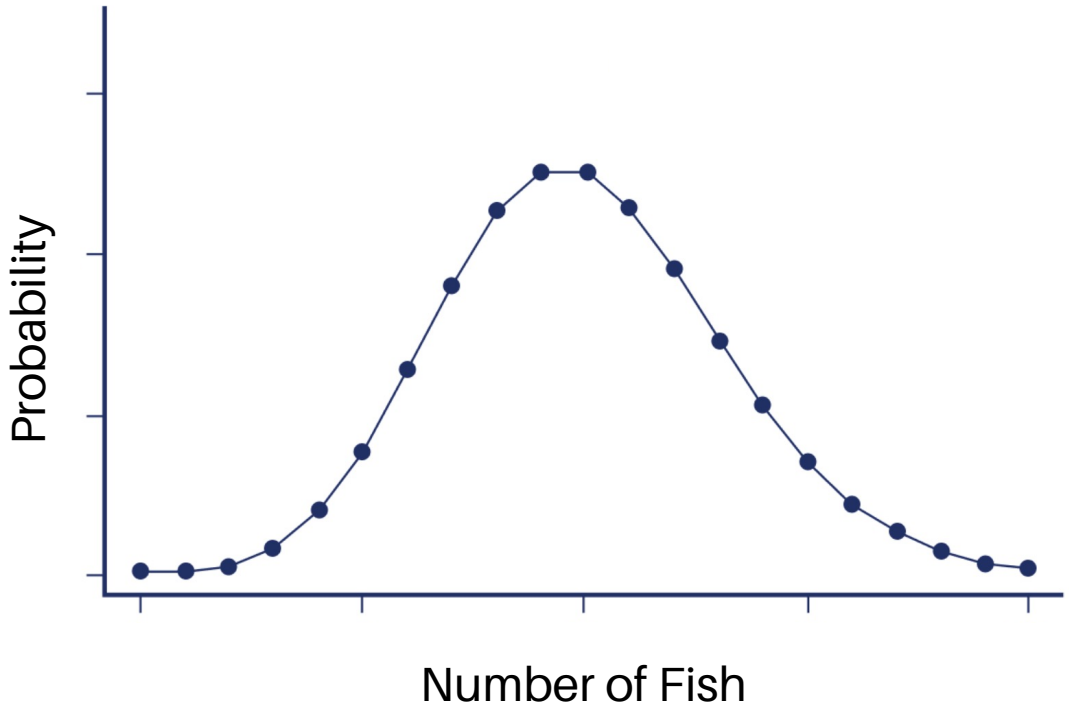
How likely is my dad to catch [#] of fish on any one trip?

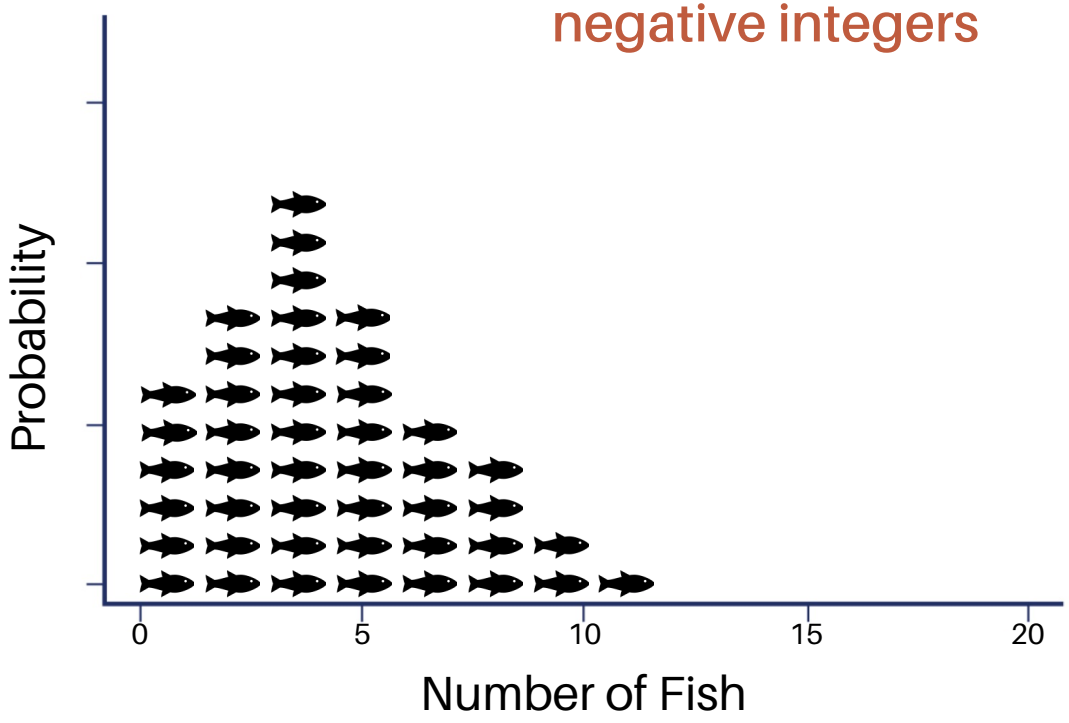




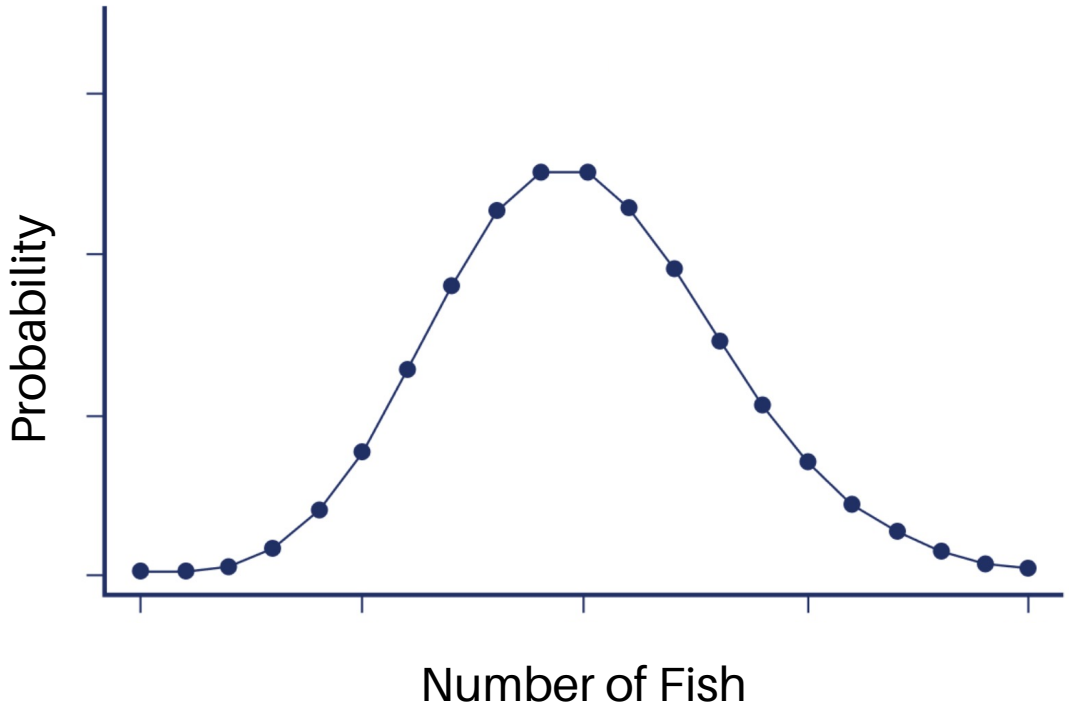
How likely is my dad to catch [#] of fish on any one trip?

(1) Counts can only be non-negative integers



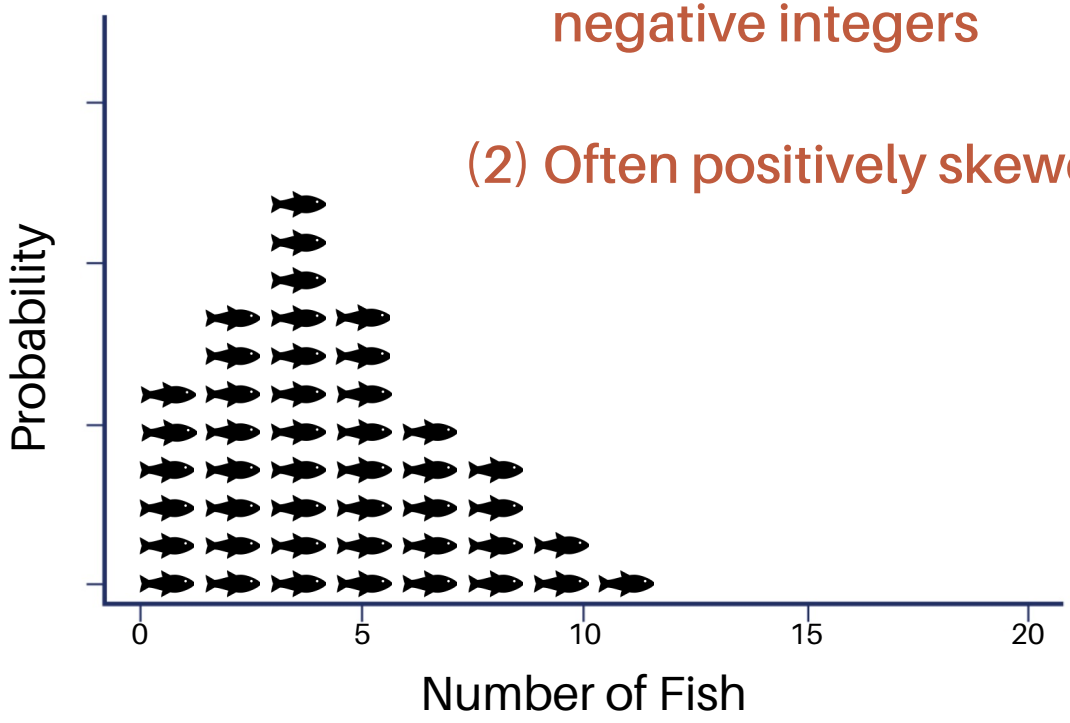


How likely is my dad to catch [#] of fish on any one trip?

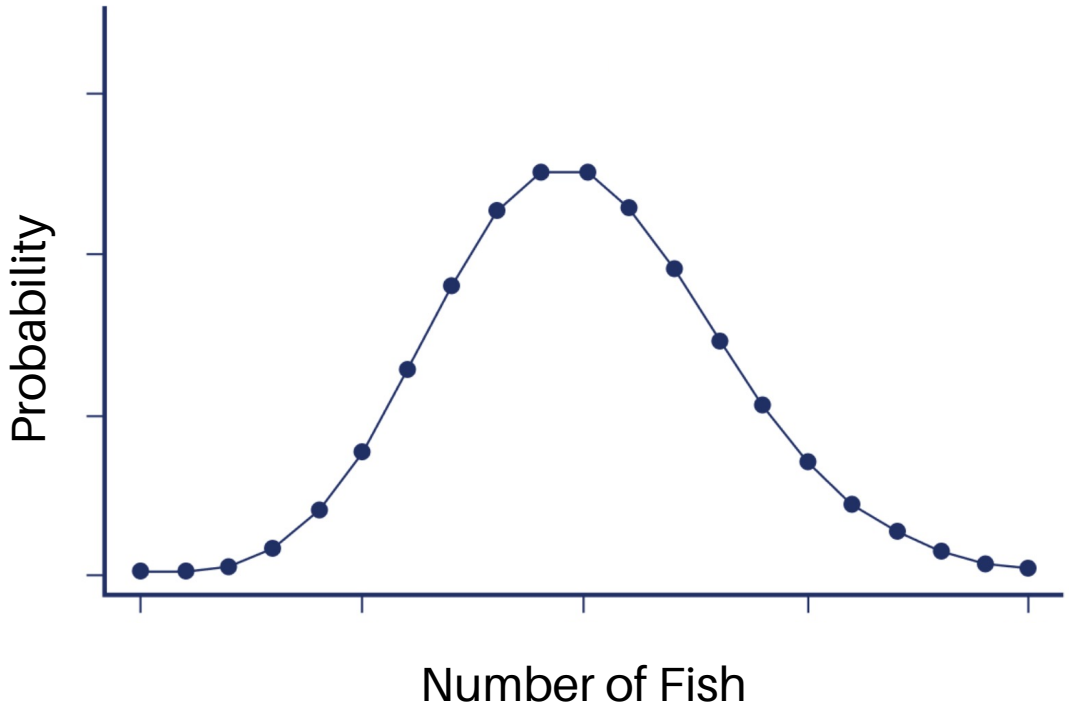


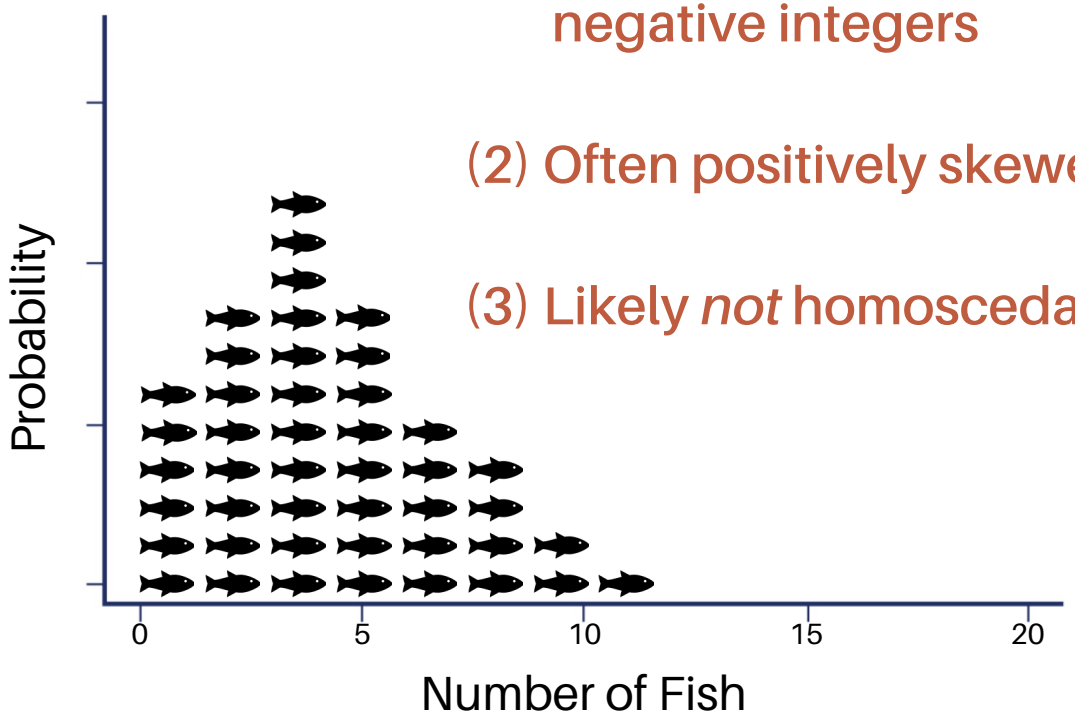
(1) Counts can only be non-negative integers

(2) Often positively skewed



How likely is my dad to catch [#] of fish on any one trip?



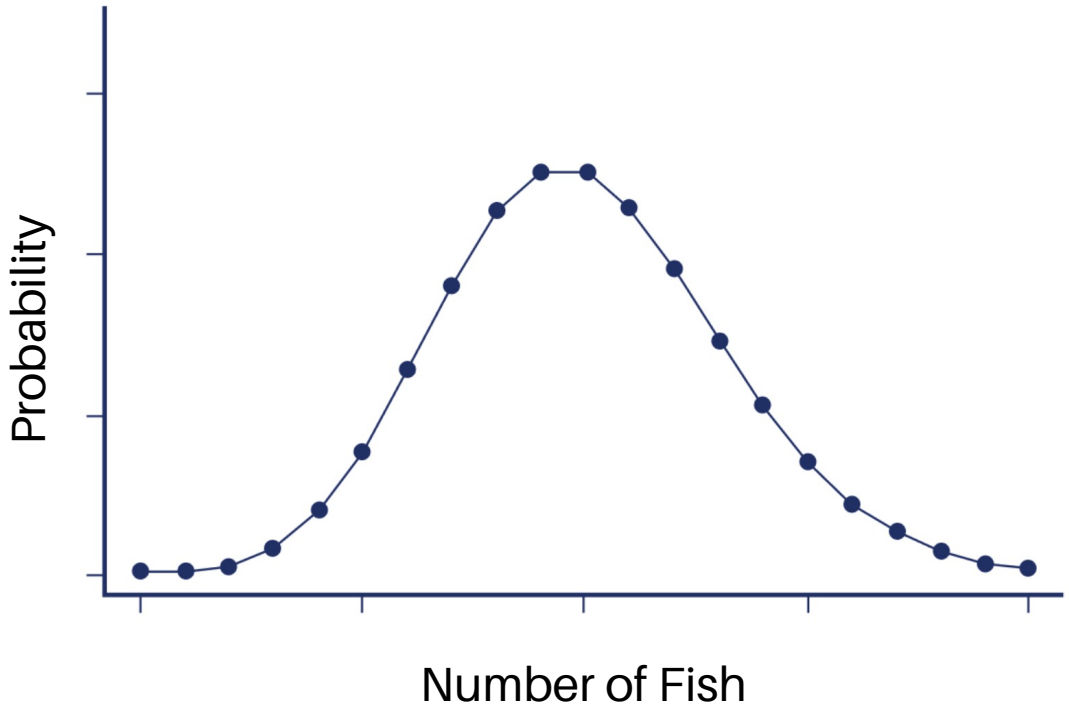


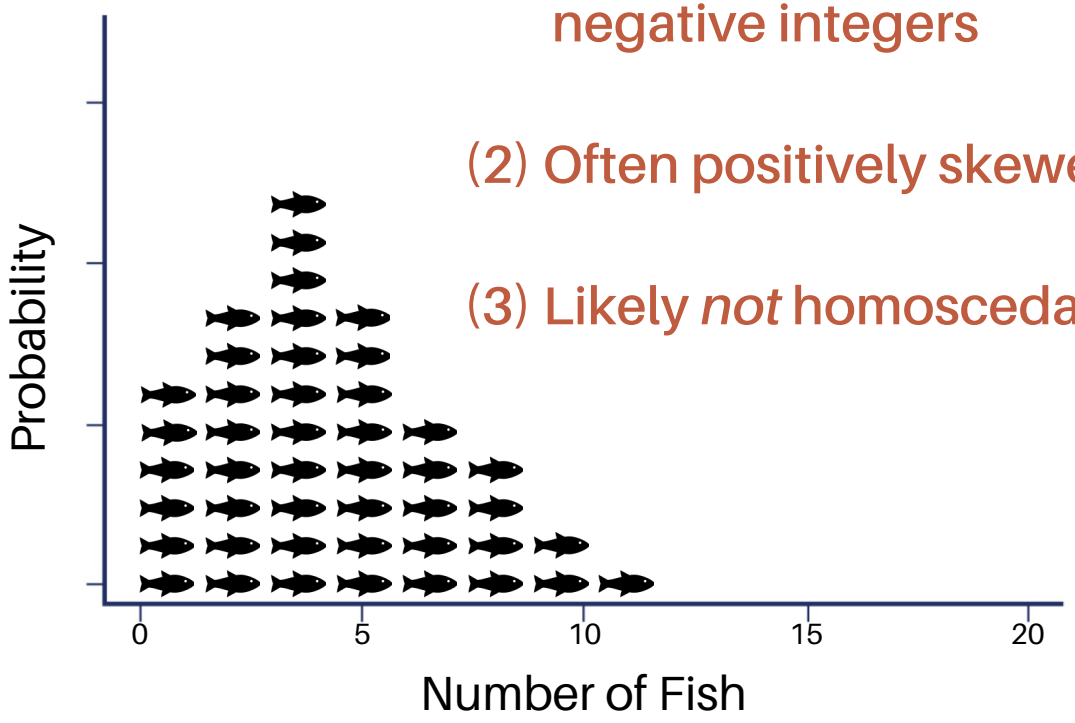
(1) Counts can only be non-negative integers

(2) Often positively skewed

(3) Likely *not* homoscedastic

How likely is my dad to catch [#] of fish on any one trip?





(1) Counts can only be non-negative integers

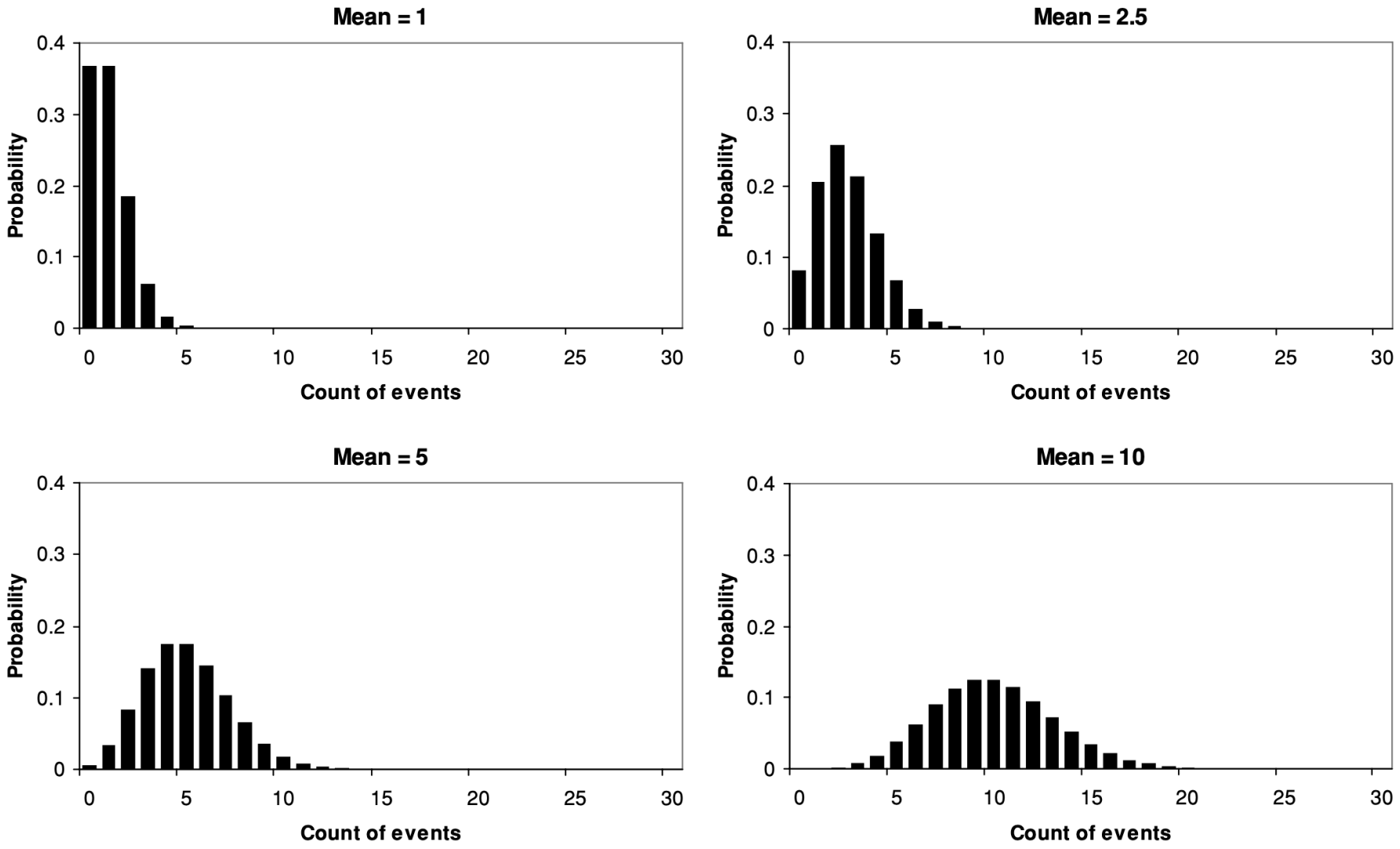
(2) Often positively skewed

(3) Likely *not* homoscedastic

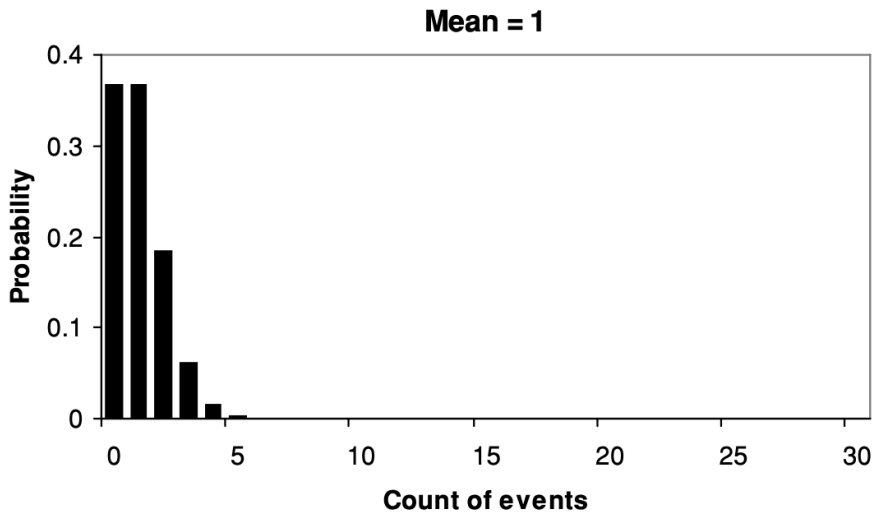
As a general rule, if the mean > 10, the poisson distribution approaches a normal distribution.

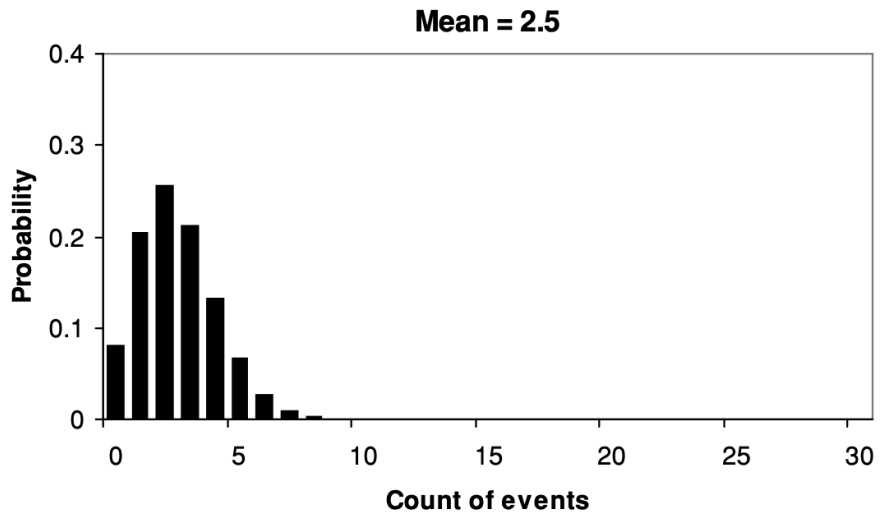
Poisson distributions

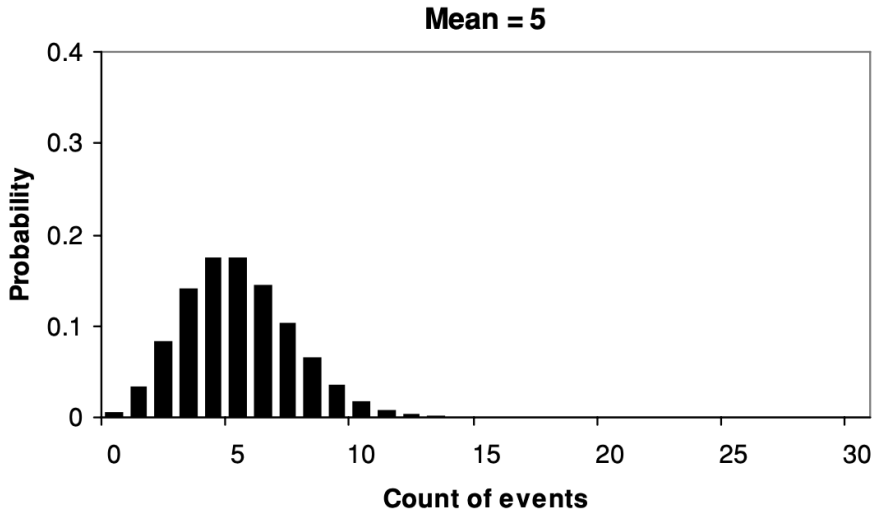
Poisson distributions

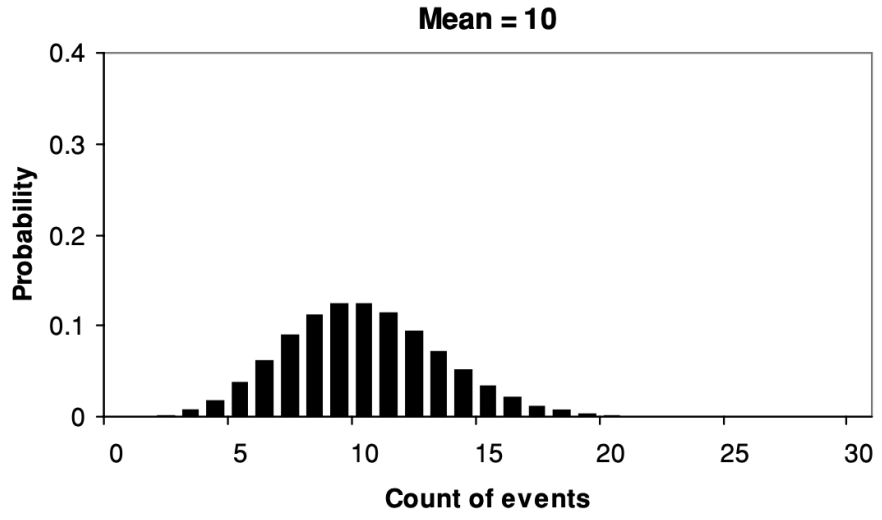


Poisson distributions



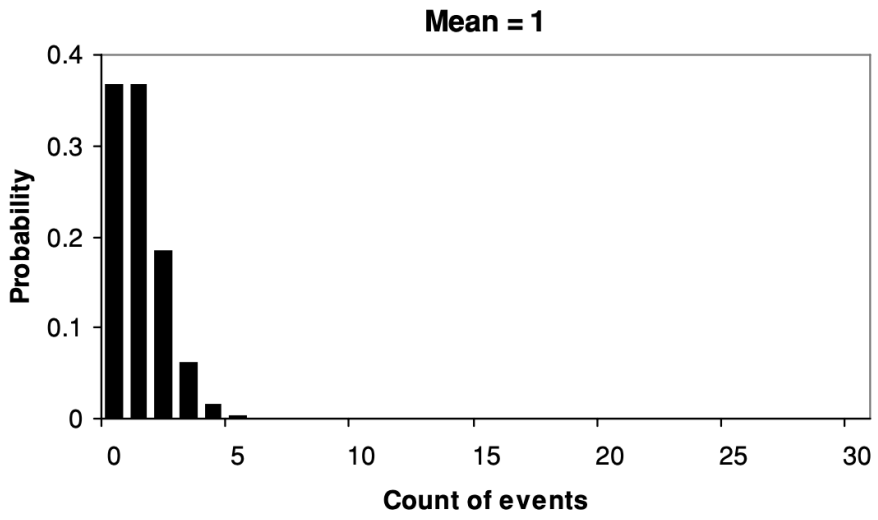


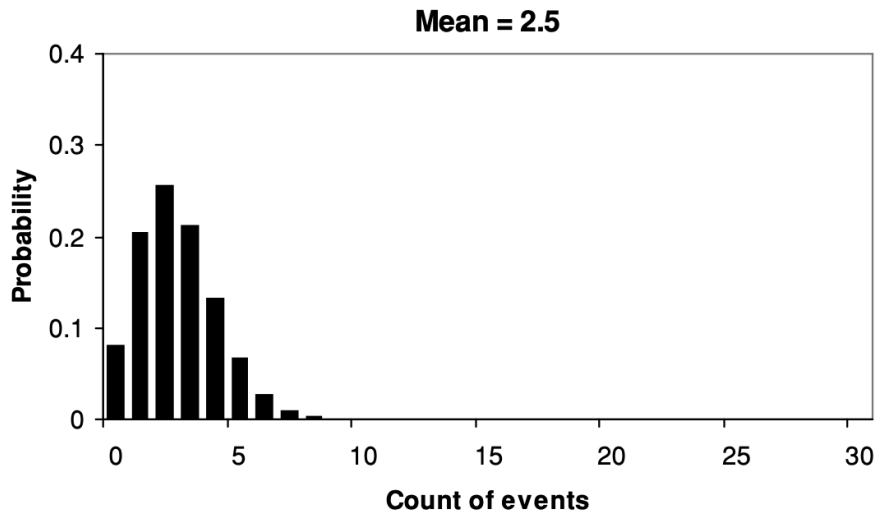


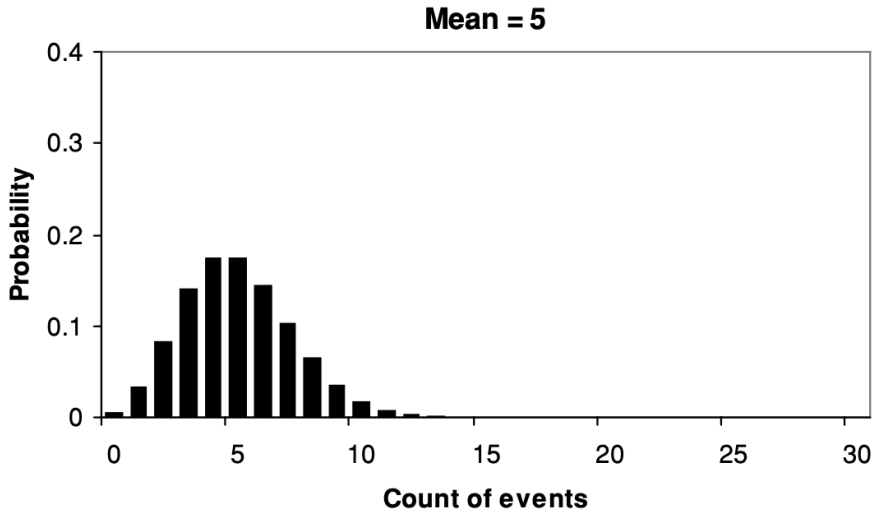


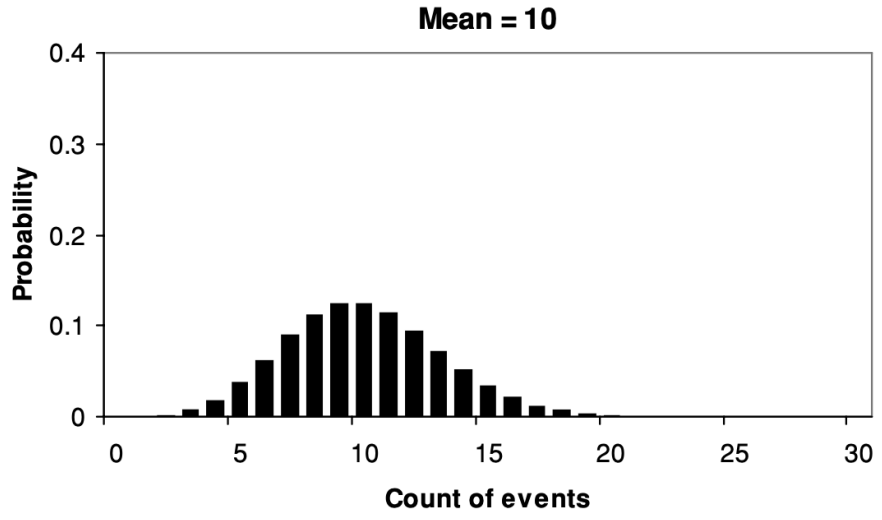
One parameter, μ, defines BOTH the mean and the variance

Poisson distributions









One parameter, μ, defines BOTH the mean and the variance

All observations are independent events

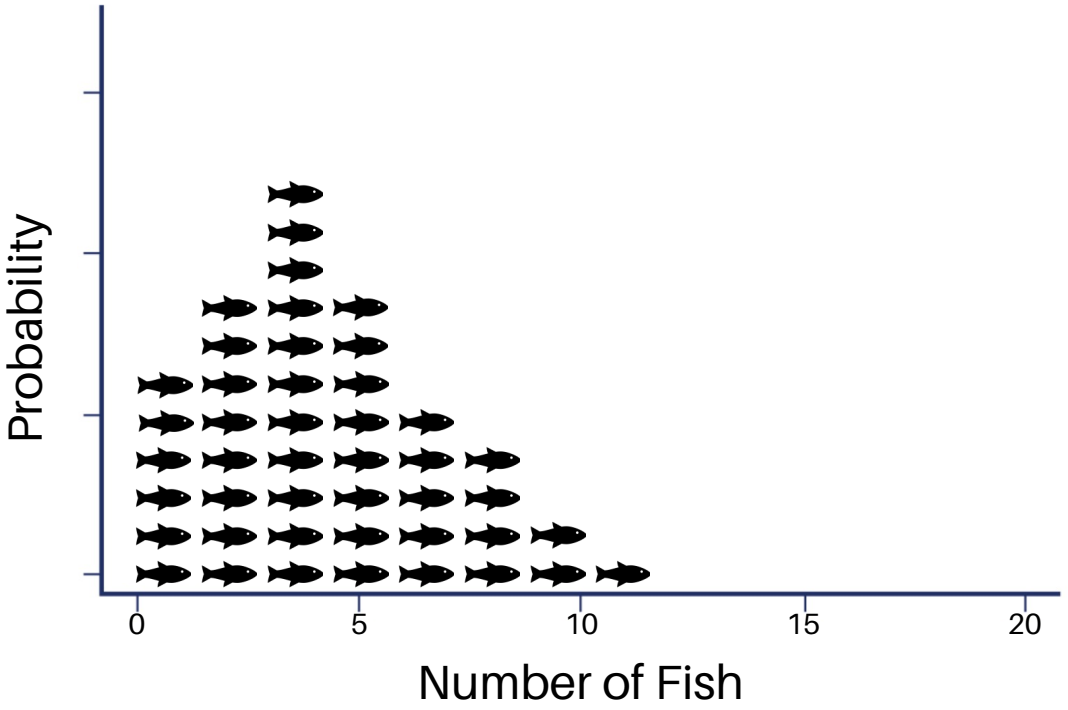
Problems?

Problems?: Overdispersion

When the variance is larger than the mean

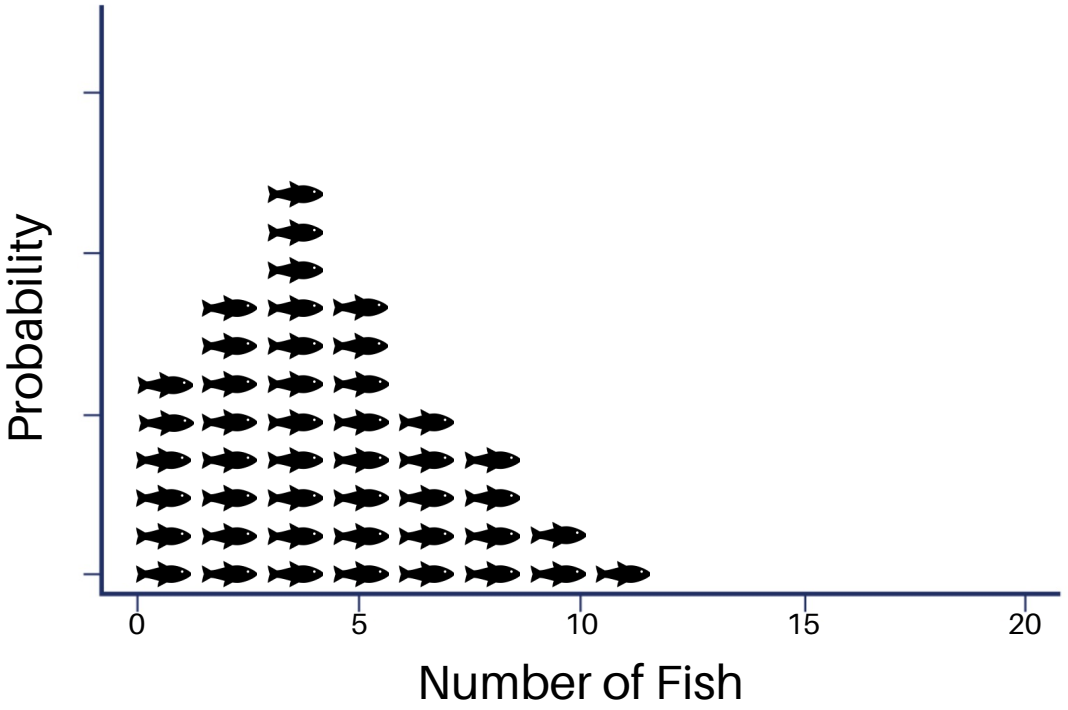
Problems?: Overdispersion

When the variance is larger than the mean



Problems?: Overdispersion

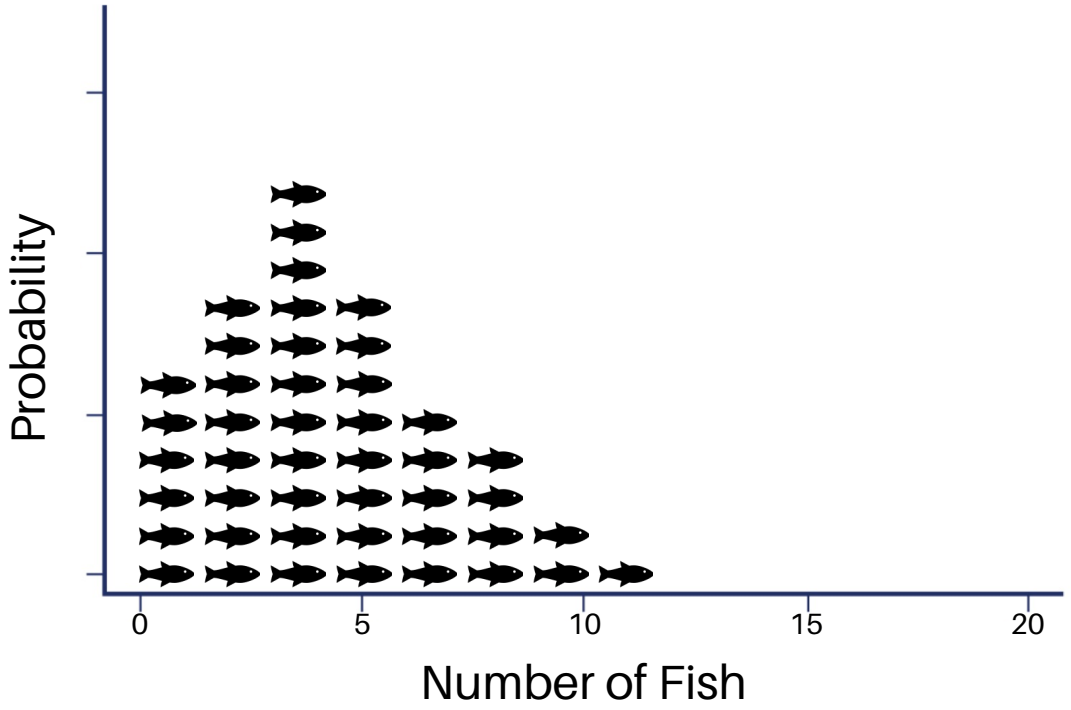
When the variance is larger than the mean

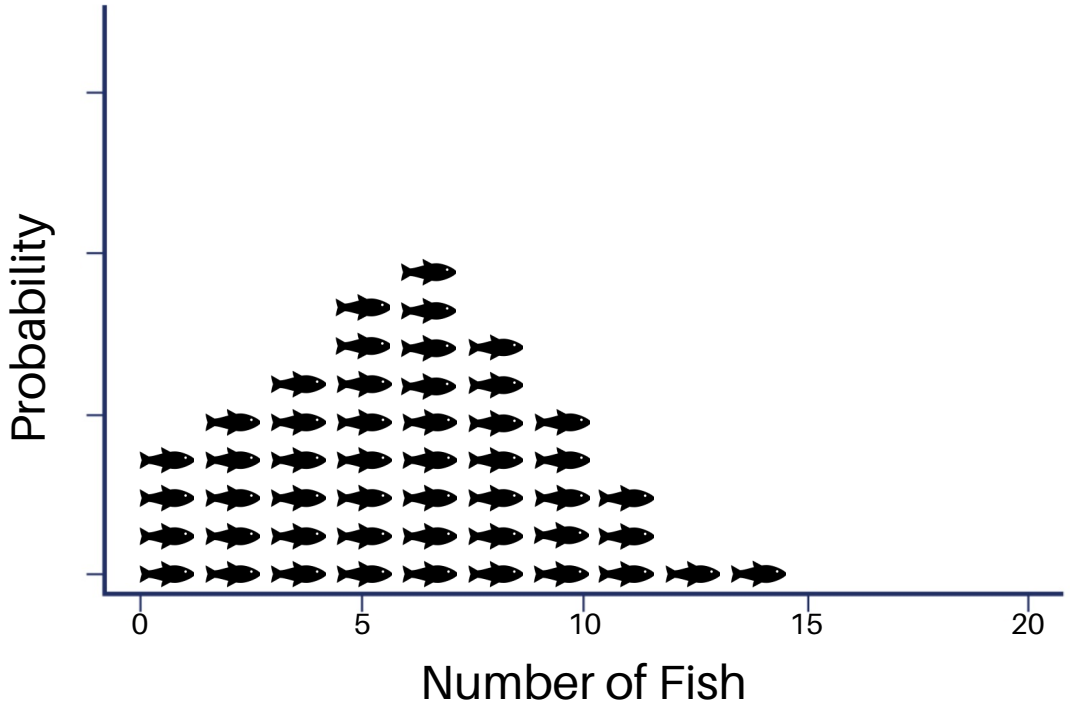


Can occur when there are differences (e.g., a predictor that's been omitted) not accounted for by the model

Problems?: Overdispersion

When the variance is larger than the mean

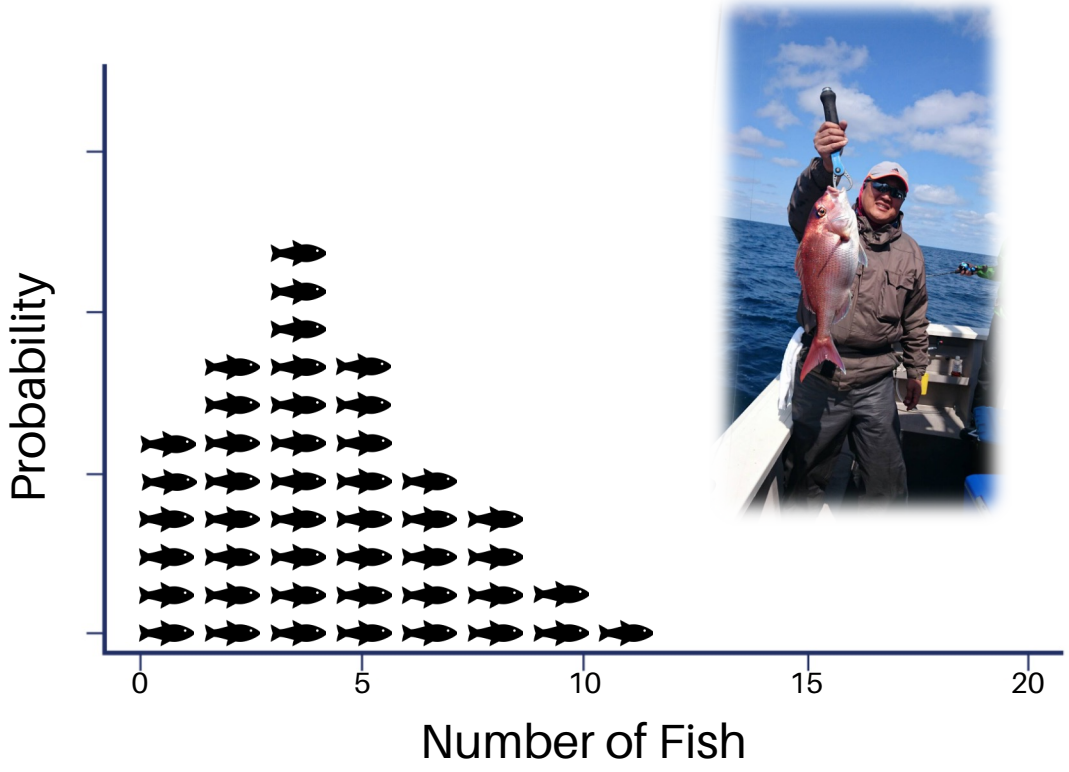


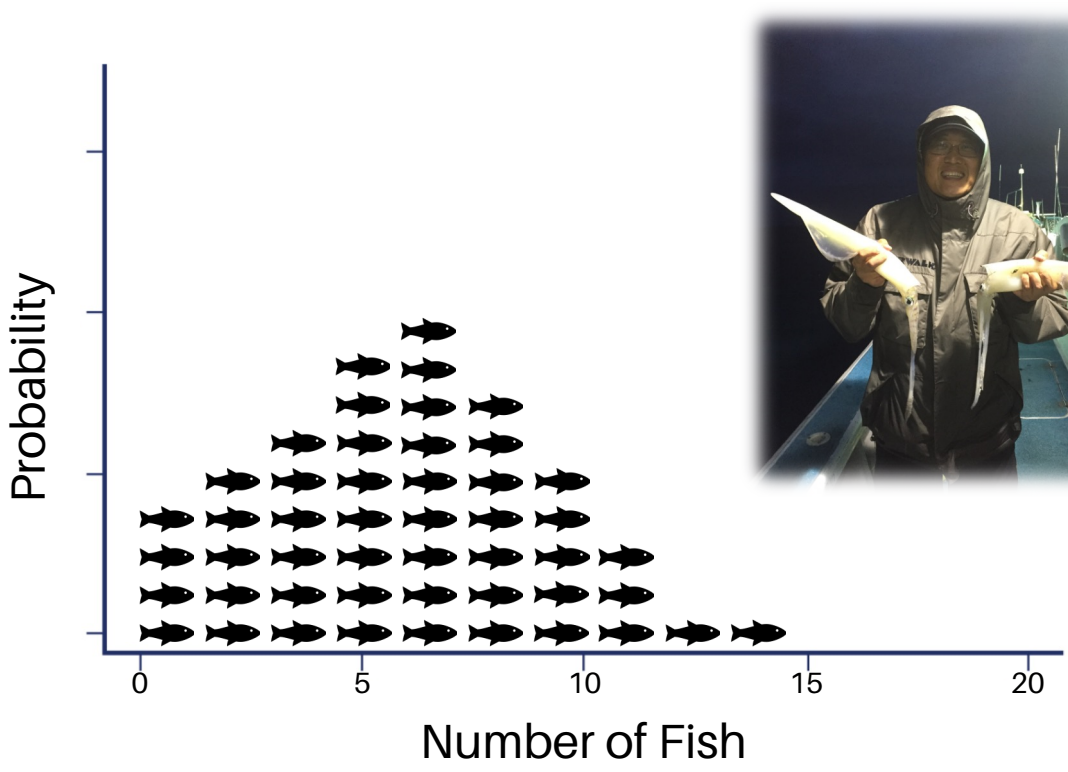


Can occur when there are differences (e.g., a predictor that's been omitted) not accounted for by the model

Problems?: Overdispersion

When the variance is larger than the mean

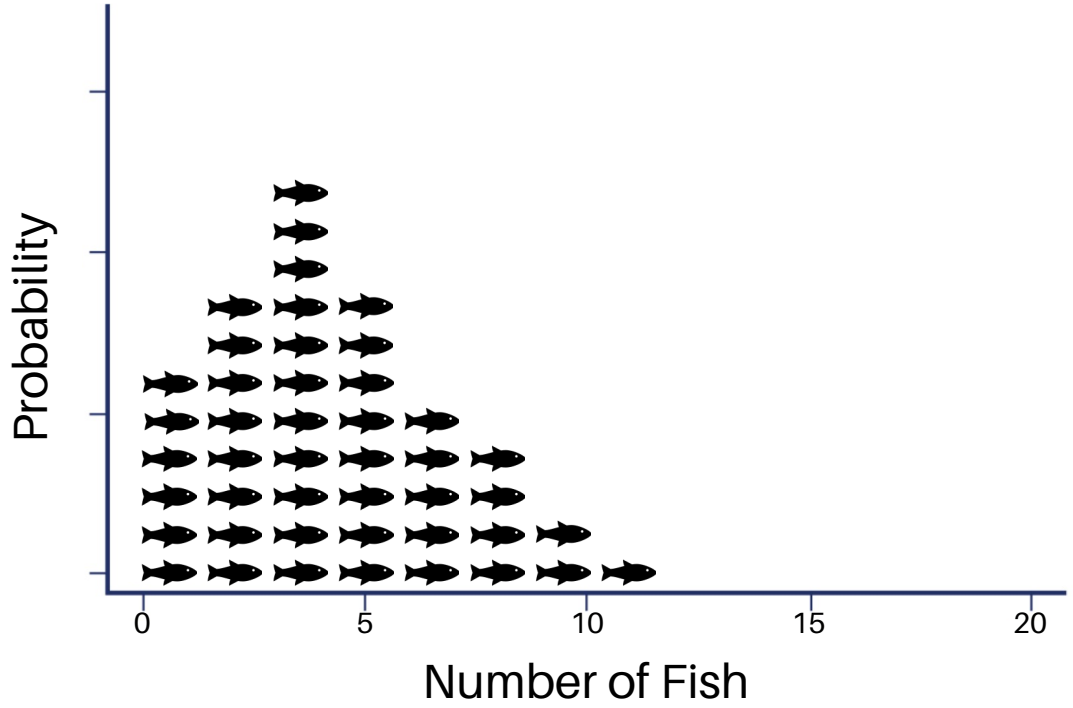




Can occur when there are differences (e.g., a predictor that's been omitted) not accounted for by the model

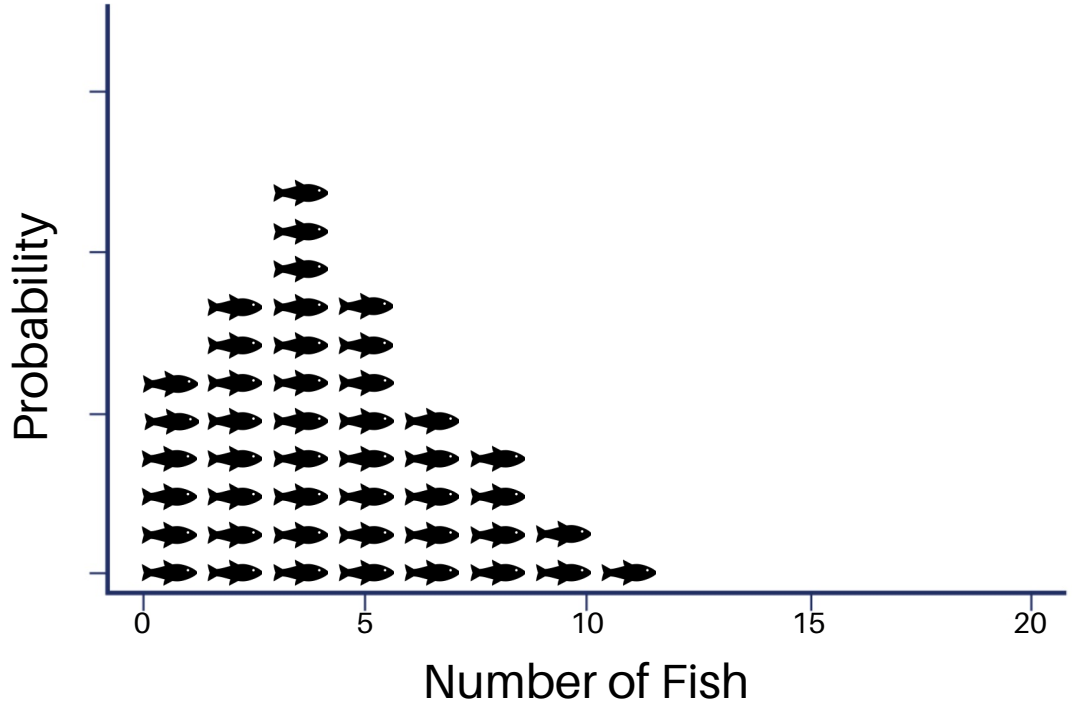
Problems?: Overdispersion

When the variance is larger than the mean



Problems?: Overdispersion

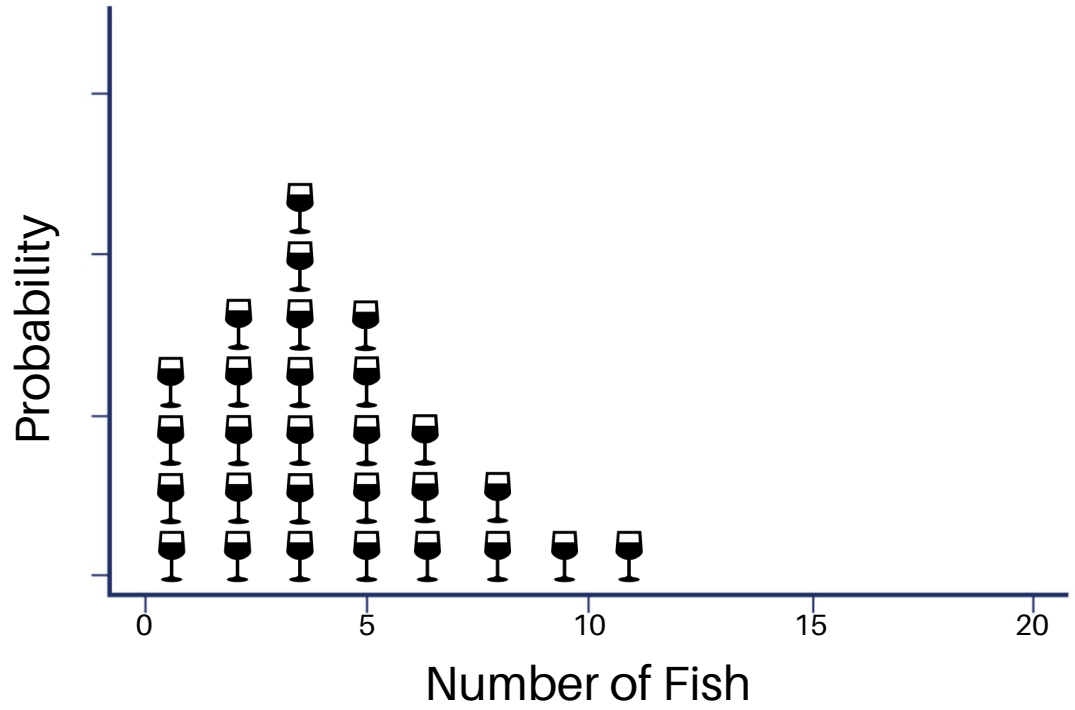
When the variance is larger than the mean



Can occur when events are not actually independent

Problems?: Overdispersion

When the variance is larger than the mean



Can occur when events are not actually independent

Problems?: Overdispersion

When the variance is larger than the mean

Can occur when there are differences (e.g., a predictor that's been omitted) not accounted for by the model

Can occur when events are not actually independent

(1) Run an overdispersed model with a second parameter that defines the variance.

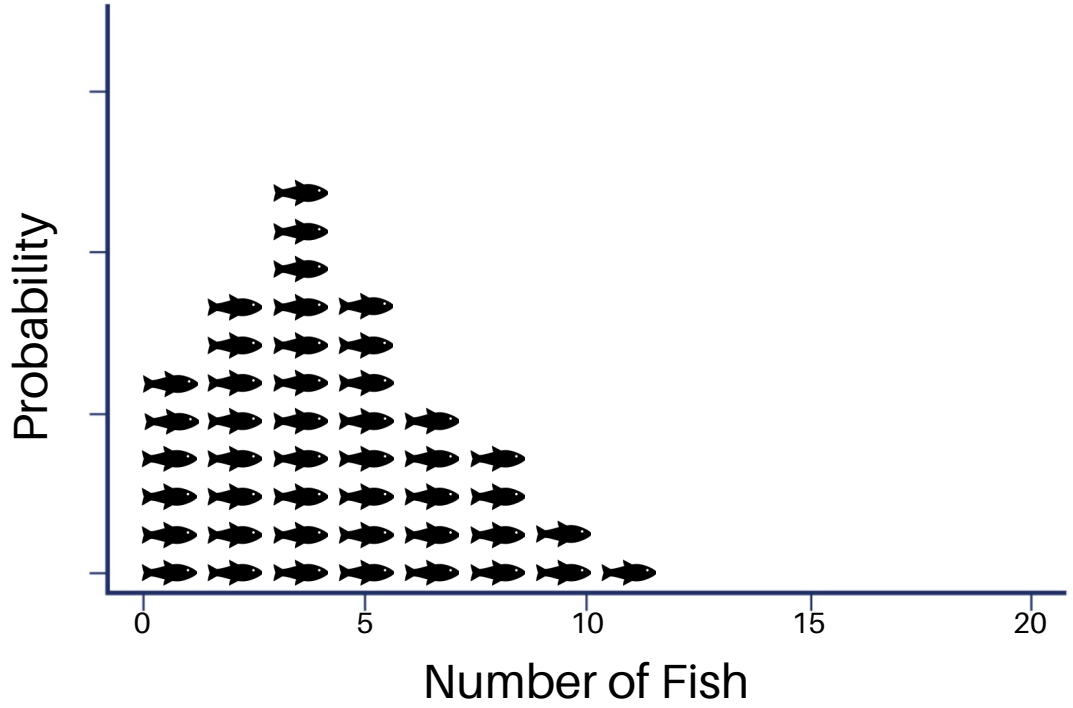
(2) Run a negative binomial model instead, which assumes that there will be unexplained variability between individuals with the same predicted value.

Problems?: Zero inflated models

When structural zeros artificially increase positive skew

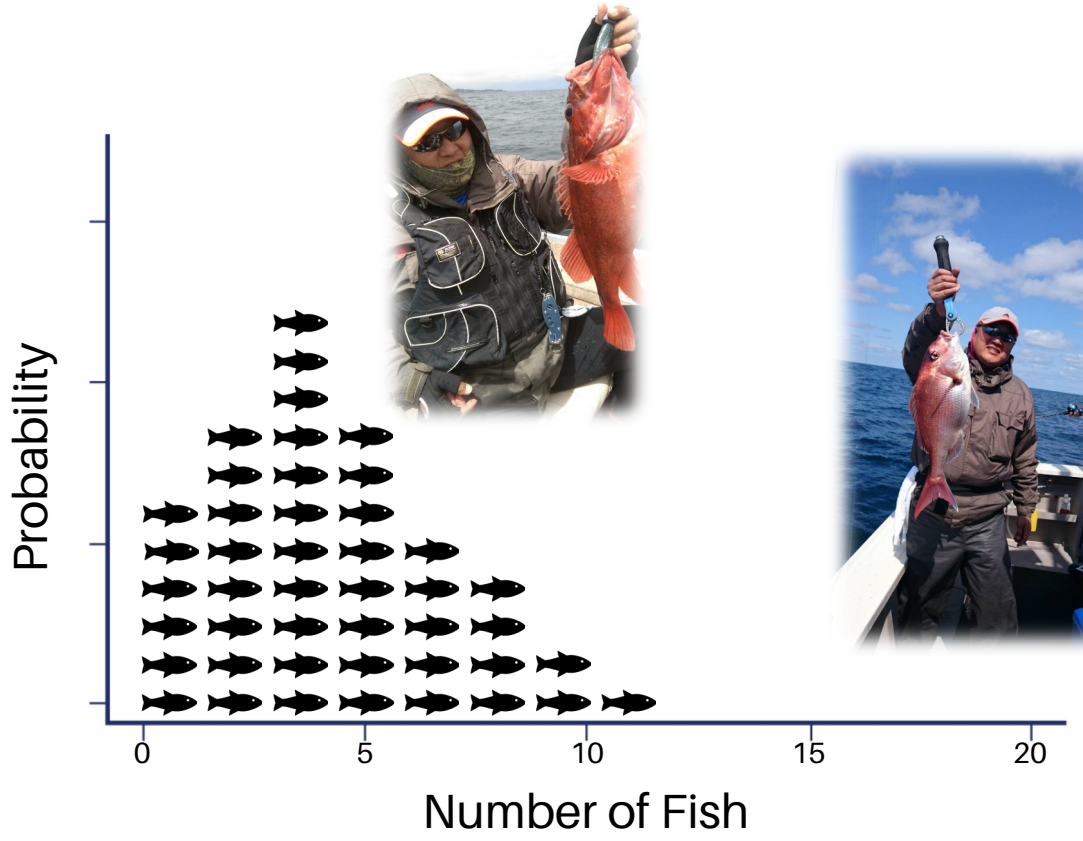
Problems?: Zero inflated models

When structural zeros artificially increase positive skew



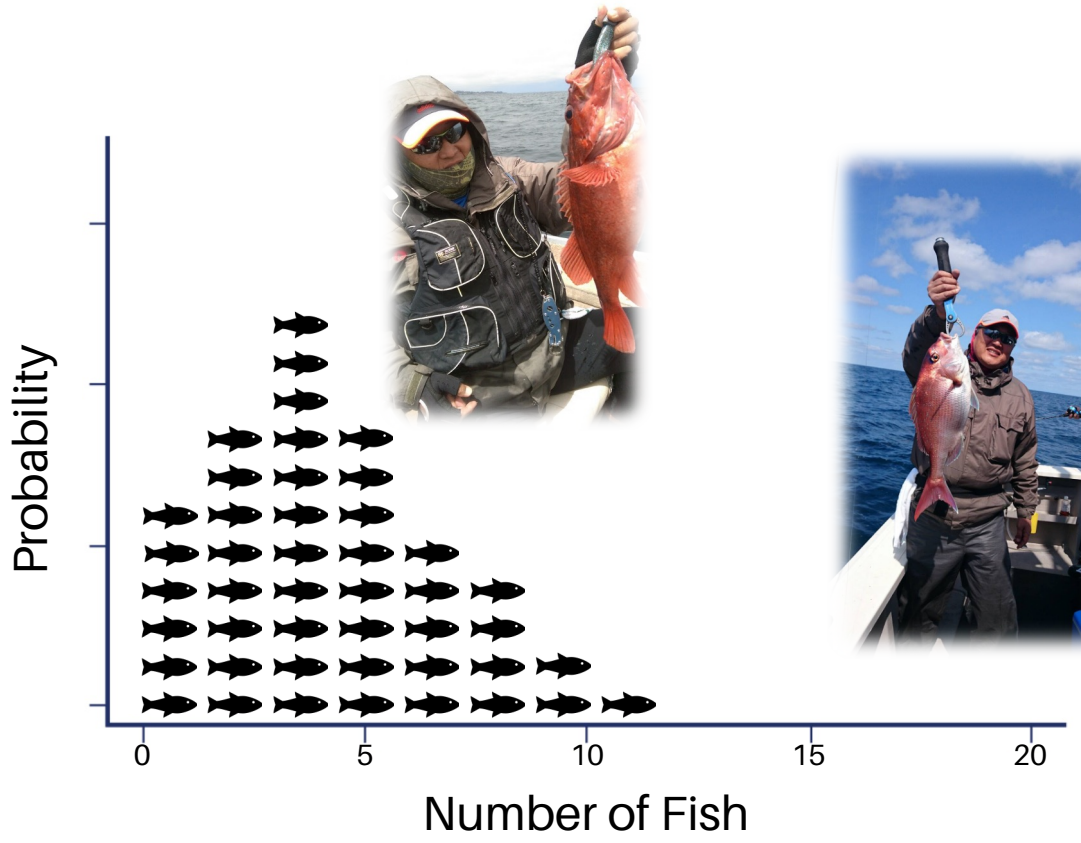
Problems?: Zero inflated models

When structural zeros artificially increase positive skew



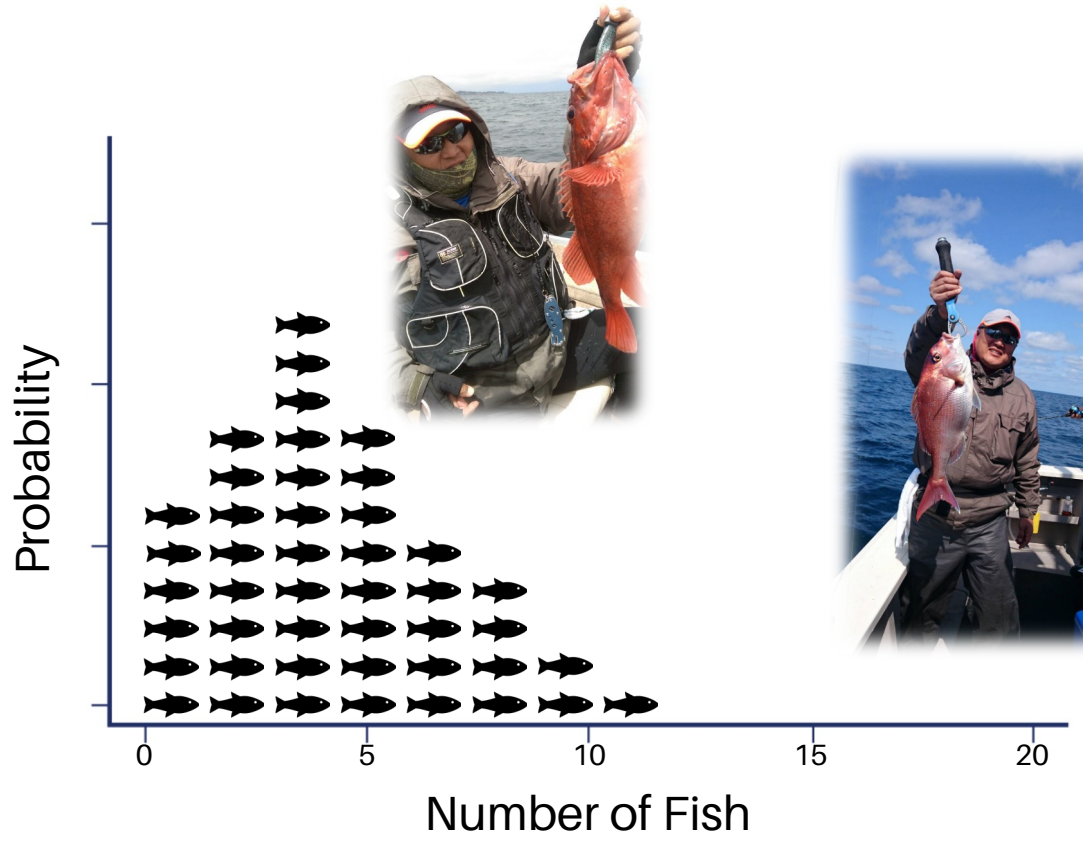
Problems?: Zero inflated models

When structural zeros artificially increase positive skew



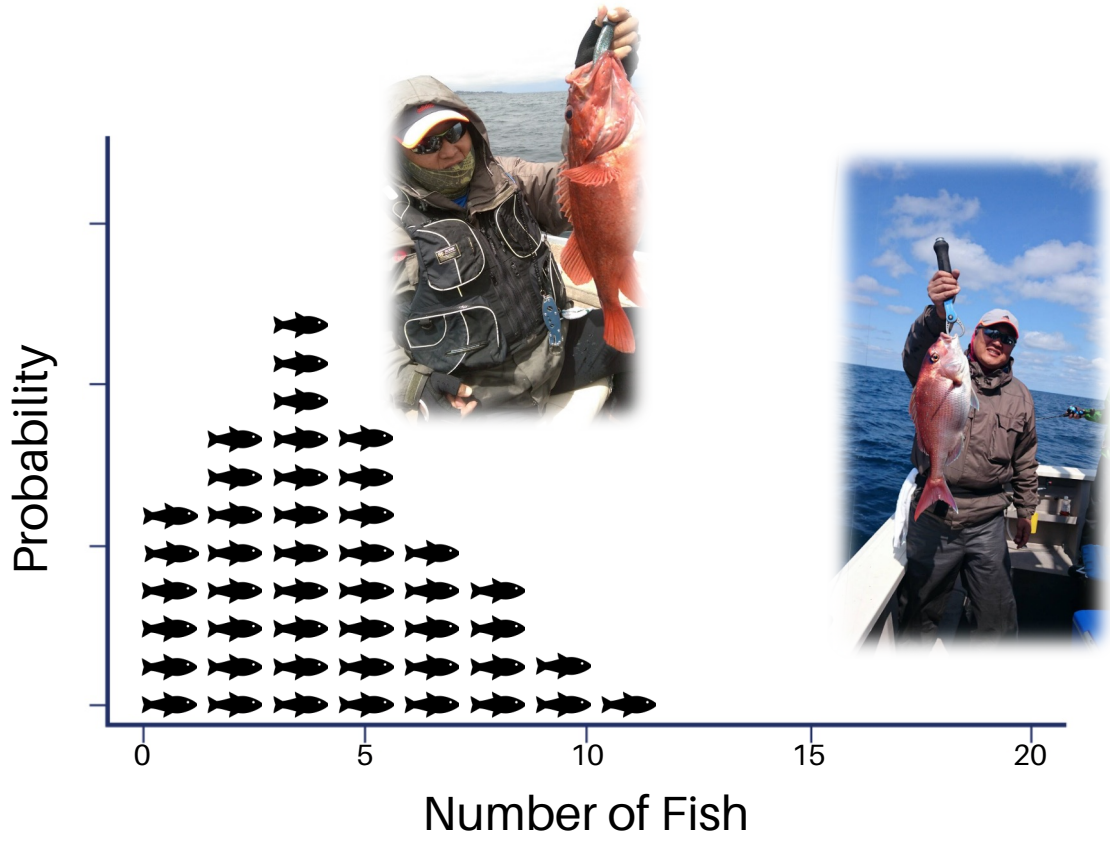
Problems?: Zero inflated models

When structural zeros artificially increase positive skew



Problems?: Zero inflated models

When structural zeros artificially increase positive skew



Best solution: Vet your sample beforehand

Other questions

Other questions

In practice, is overdispersion ever not a problem? Why ever run a poisson regression over a negative binomial regression?

In reference to a previous conversation: To what extent are these concerns mitigated by taking a model comparison approach?