Modeling number of seeds in a pod

- **Model: mathematical representation of reality**

- Ecological models: two types
  - Explanatory (evolutionary, survival of the fittest)
  - Descriptive (summarizing)

- Clutch size in birds
  - Vulture –1-2, Eagle – 2-3, Myna – 4-5
  - (Contrast: Fish – thousands)

- Aim : maximize # viable offsprings
  - Too many offspring- feeding inadequate (parental capacity)

An explanatory model for clutch size in birds

- C : clutch size
- p: prob of survival of an offspring $=1- \gamma C$
- γ : related to parental capacity
- X: # surviving offspring

- Objective function E(X)
- $E(X) = C(1- \gamma C)$
- E(X) : maximum at $C = 1/(2\gamma)$

- Parental capacity decides optimum clutch size
- Model prediction:
    - Capacity very low : **clutch size zero – avoid reproduction**
        - Bad season, shortage of time

A descriptive model for plants

- Clutch size in plants: number of seeds in a pod/fruit
  - Value varies
    - From species to species
    - From pod to pod in the same species

Species **(ovule #)**		Number of seeds											
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Caesalpinia decapetala (C.d.)	**8**	1	1	0	2	8	16	32	7	-	-	-	-
Malletia ovalifolia(M.o.)	**5**	41	41	12	1	0	-	-	-	-	-	-	-
Lablab niger (L.n.)	**5**	1	4	22	41	1	-	-	-	-	-	-	-
Albezzia lebbek (A.l.)	**12**	4	5	11	9	17	19	16	17	20	15	9	1

- C.d. and L.n. : negatively skewed
- M.o. : positively skewed
- A.l. : multimodal
- How to summarize this variety of situations?

Simplest model

- Ovule ~ egg
- Pollen ~ sperm
- On fertilization : one seed per ovule
- Assumption :
  - Each ovule a Bernoulli trial
  - Ovule number fixed for a species :n
  - no shortage of pollen
- If fertilized and seed formed : success, otherwise failure
- Each pod : a binomial experiment (n,p)
- p = prob. of success

- Pods with zero seeds : not observable because pod drops off
- Zero truncated binomial

- $P(X= r) = {}^{n}C_{r}\, p^{r}\, (1-p)^{n-r} / [1-(1-p)^{n}]$

- How good is this model?

Fitting zero truncated binomial model

How to estimate p? Use moment estimator i .e. solve the eqn.
Average x = $n*p/[1-(1-p)^n$

Species	**C.d.**	**M.o.**	**L.n.**	**A.l.**
Est(p)	0.80	0.27	0.71	0.57
Ovule #	8	5	5	12
Chisquare	3.85	2.83	23.05	95.59
d.f.	3	2	3	7

- C.d., L.n. negatively skewed: p>0.5
- M.o. positively skewed : p<0.5
- Model acceptable for C.d. and M.o. (and many more)
- For L.n. and A.l.: modification needed
    - Relax some assumption

Truncated binomial model: mixture over n

- Ovule # : not quite fixed for a species, varies a bit

- We used largest **observed** value

Lablab niger		Albezzia lebbek	
#ovules	Proportion of flowers	#ovules	Proportion of flowers
3	3/16	11	2/3
4	3/4	12	1/3
5	1/16	-	-

$$\mathbf{P(X=r) = \Sigma\alpha_i\, B(\, n_i\, ,p)}$$

Fitting mixture over n

Species	L.n.	A.l.
Est(p)	0.92	0.60
Ovule #	3,4,5	11,12
Chisquare	1.52	89.70
d.f.	2	5

- Mixture model adequate for Lablab niger
- Clearly inadequate for Albezzia lebbek

- Further modification needed
- Which assumption to relax?
- Adequacy of pollen supply
- In A.l. pollen grains come in packs of 4

Pollen limiting model

- If only one pack of pollen received (prob. λ_4): B(4,p)
- If two packs received (prob. λ_8): B(8,p)
- If three or more received (prob. λ_{12}):

  Pollen are not limiting, ovules are.
  - 2/3 B(11,p) + 1/3 B(12,p)

Model:

$P(X=r) = \lambda_4 B(4,p) + \lambda_8 B(8,p) + (1-\lambda_4-\lambda_8)[2/3\ B(11,p) + 1/3\ B(12,p)]$

Estimates: $p = 0.76$, $\lambda_4 = 0.18$, $\lambda_8 = 0.32$

Chisquare value 3.35 with 6 d.f.

Satisfactory fit

Butterfly intervention model

- Attempt at physical ***interpretation*** of seed number distribution
  - Negatively skewed: Avoidance by parent of few-seeded fruits
    - Economical use of packaging material (parent interest)
  - Positively skewed : A few dominant seeds causing abortion of others (offspring interest)
  - Unimodal symmetric : balance between parent interest and offspring interest.
    - What about bimodal distribution?
    - Does it exist?

Distribution of seeds /pod in Caeslapinia pulcherima

Seed #	1	2	3	4	5	6	7	8	9
Frequency	10	25	27	9	12	17	31	12	1

Binomial mixture over p

- Model :
  - $P(X=r) = {}^{n}C_{r} \{ \lambda p_1^{r}(1-p_1)^{n-r} + (1-\lambda) p_2^{r}(1-p_2)^{n-r} \}$
    - r = 1,2,…n
    - n: # ovlules
    - r = 0 (pods with zero seeds not available)
    - p1,p2: Prob. of ovule fertilizing, maturing to seed

Fitting the model (zero truncated)

Est(λ)	Est(p_1)	Est(p_2)	Chisq.	d.f.
.51	.28	.73	14.32	5

Interpretation

- p1,p2: Prob. of ovule fertilizing, maturing to seed
  - Estimates : 0.28, 0.73
  - Why?
  - Species butterfly pollinated
- Assumption:
  - Pollen limiting situation
    - Low p: low availability
- Speculation:
  - Butterfly inserts proboscis into flower: large clump of pollen
  - Merely flutters wings while on flower : few pollen
  - Testable hypothesis

Conclusion

- A simple model could describe a biological phenomenon in many species

- Relaxing assumptions increases applicability of a model

- Model sometimes suggests directions for new observations
  - In case of A.l. empirical verification of λ_4, λ_8 necessary
  - In case of C.p. verification of λ

λ=0.51:Both situations equally frequent

Home work:
Collect data on seed count for a species of your choice
Try fitting one of the models described