Review of Probability Theory

- Measurements of random events may be made in either continuous or discrete time.
- Examples are, height, number of heads when a coin is tossed 10 times.
- Random variables are characterized by a probability density function (probability mass function for discrete random variables), $g(x)$, which have the property that $\int_{-\infty}^{\infty} g(x)dx = 1$, or $\sum_{i=a}^{i=b} g(x_i) = 1$, where x_i takes on values, x_a, …, x_b.
- Examples: uniform distribution on (0,2), $g(x)$=0.5,
  $\int_0^2 0.5dx = 0.5x \Big[_0^2 = 0.5(2-0) = 1$

Review of Probability Theory (cont.)

- Binomial distribution, $g(x) = \binom{n}{x} p^x (1-p)^{n-x}$, where

$$\binom{n}{x} = \frac{n!}{x!\,(n-x)!}$$

- mean=np and variance=np(1-p)
- Example, with n=2, p=0.6,

$$\sum_{x=0}^{x=2} \binom{2}{x} 0.6^x (1-0.6)^{2-x} = 1 \cdot 0.6^0 \cdot 0.4^2 + 2 \cdot 0.6^1 \cdot 0.4^1 + 1 \cdot 0.6^2 \cdot 0.4^0$$
$$= 0.16 + 0.48 + 0.36 = 1$$

Expectation

- The expected value of a function, *f*(*x*), of a random variable is, E[*f*(*x*)]=$\int_{-\infty}^{\infty} f(x)g(x)dx$ or E[*f*(*x*)]= $\sum_{i=a}^{i=b} f(x_i)g(x_i)$
- Example, let *f*(*x*)=*x*, then using the previous examples,
  Uniform: $\int_0^2 0.5xdx = 0.5 \cdot 0.5x^2 \Big[_0^2 = 0.25(4-0) = 1$
  Binomial: $0.16 \cdot 0 + 0.48 \cdot 1 + 0.36 \cdot 2 = 0 + 0.48 + 0.72 = 1.2$
- The mean of a random variable can be defined as, E[*x*], and the variance, E[(*x*-*E*(*x*))^2].
- Cov(xy)= E{[x-E(x)][y-E(y)]}

Expectation and Variance

- $E[cx]=cE[x]$, where c is a constant
- $E[k+cx]=k+cE[x]$, where k and c are constants
- $Var(c)=0$
- $Var(cx)=c^2Var(x)$ -> prove this by using the deinfinition of variance, $E[(x-E(x))^2]$
- Show that, $Var(x)= E(x^2)-E(x)^2$ and $Cov(xy)=E(xy)-E(x)E(y)$
- $Var(x+y)=Var(x)+Var(y)+2Cov(xy)$ -> prove by using the definition of variance

Conditional Expectation and Probability

- For discrete random variable
- E[Y]=$\sum_{over\ x} E[Y|X = x]g(x) = E[E[Y|X = x]]$
- If X and Y are independent, E[Y|X=x]=E[Y]
- $E[f(X,Y)|X = x] = E[f(x,Y)|X = x]$
- Var[Y]=E[Var(Y|X)]+Var[E(Y|X)]

Conditional Mean and Variance: Example

- Example: Suppose we have many vials with fruit flies. The number of females in a vial (*N*) is a random variable with a binomial distribution, ~$B(\tilde{N},p)$, where $\tilde{N}$ is the total number (constant) of flies placed in a vial. Suppose the number of eggs laid by each female (*X*) has a Poisson distribution, with parameter λ (mean and variance $=\lambda$). Then the total number of eggs laid in a vial is, $Y=X_1+X_2+\ldots+X_N$.
- $E[Y|N=n]=nE[X]$-> $E[Y] = E[X]\sum_{j=0}^{\tilde{N}} j\binom{\tilde{N}}{j}p^j(1-p)^{\tilde{N}-j} = E[X]\tilde{N}p = \lambda\tilde{N}p$

Conditional Mean and Variance: Example (cont.)

- $\text{Var}(Y)=\text{E}[\text{Var}(Y)|N=n]+\text{Var}(\text{E}[Y|N=n])=\text{E}[n\text{Var}(X)]+\text{Var}(n\text{E}(X))$,
  since $\text{Var}(Y)|N=n = \text{Var}[X_1+X_2+\ldots+X_n]=n\text{Var}(X)$ and
  $\text{E}[Y|N=n]= \text{E}[X_1+X_2+\ldots+X_n]=n\text{E}(X)$

  $\text{Var}(Y)=\text{E}(N)\text{Var}(X)+\text{E}(X)^2\text{Var}(N)$,
  since Var(X) and E(X) are constants

  $\text{Var}(Y)=\tilde{N}p\lambda + \lambda^2 p(1-p)\tilde{N}$

- If $\tilde{N}=100$, $p=0.5$, $\lambda=10$, then $\text{E}[Y]=500$, and $\text{Var}(Y)=3{,}000$
  an approximate 95% confidence interval is $2 \cdot \sqrt{3000} = 110$

Taylor Series

- The value of a function near a point x^* can be approximated with the Taylor Series.
- For the function $f(x)$ let $f^{(n)}(x^*)$ be the *nth* derivative evaluated at the point x^*, then

$$f(x) = f(x^*) + \frac{(x - x^*)}{1!} f^{(1)}(x^*) + \cdots + \frac{(x - x^*)^n}{n!} f^{(n)}(x^*)$$

- Example: exponential function at $x^*=0$,

$$e^x \cong 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}$$

- If we just use the first two terms, $e^{0.1} \cong 1.1$, while the true answer is 1.105.
- If x is far from x^* then the approximation will be poor. So $e^5 \cong 6$, while the true value is 148.

Delta Method

- The delta method is a way of approximating the variance of complicated functions.
- M, which is a complicated function of k parameters, c_1, c_2, …, c_k, e. g. $M = F(c_1,c_2,., c_k)$.
- The parameters are estimated as, $\hat{c}_1, \hat{c}_2, \dots, \hat{c}_k$. Then, $\widehat{M} \cong F(\hat{c}_1, \hat{c}_2, \dots, \hat{c}_k)$.
- The delta method uses a Taylor series expansion around E(c_i) which in practice will be estimated by $\hat{c}_i$.

$$\widehat{M} \cong F[E(c_1), E(c_2), \cdot\cdot, E(c_k)] + (c_1 - E(c_1))\frac{dF}{dc_1}|_{c_1=E(c_1)} + \cdot\cdot + (c_k - E(c_k))\frac{dF}{dc_k}|_{c_k=E(c_k)}$$

- Then noting that $Var(\hat{M}) = E\left[\left(\hat{M} - F\left[E(c_1), \cdot\cdot, E(c_k)\right]\right)^2\right]$
- $$\sum_i Var(\hat{c}_i)\left(\frac{dF}{dc_i}\right)^2 + \sum_i \sum_{j \neq i} Cov(\hat{c}_i, \hat{c}_j)\frac{dF}{dc_i}\frac{dF}{dc_j}$$

Delta method: example

- Let X and Y be independent random variable with means, μ_x and μ_y and variances σ_x^2, and σ_y^2.
- Find Var(X/Y)
- $\frac{\partial\left(\frac{X}{Y}\right)}{\partial X} = \frac{1}{Y}$, evaluated at μ_y yields $1/\mu_y$
- $\frac{\partial\left(\frac{X}{Y}\right)}{\partial Y} = -X/Y^2$, evaluated at μ_x and μ_y yields, $\frac{-\mu_x}{\mu_y^2}$
- $Var\left(\frac{X}{Y}\right) = \frac{\sigma_x^2}{\mu_y^2} + \sigma_y^2\frac{\mu_x^2}{\mu_y^4} = \frac{\mu_x^2}{\mu_y^2}\left[\frac{\sigma_x^2}{\mu_x^2} + \frac{\sigma_y^2}{\mu_y^2}\right]$

Euphydryas editha: example

- 1985, Genetics 110: 495
- Estimates of effective population size require an estimate of the variance in reproductive success
  $N_e = \frac{2(2N-1)}{\left({}^{Vm}/_{2}+3\right)}$ (assuming equal sex ratio, no pop growth, male variance =mean)
- X_i number of eggs laid in the *ith* egg mass, with mean μ_{X_i} and variance $\sigma^2_{X_i}$.
- Host plant desiccation -> $Y \begin{cases} 0\; w.p.\, \delta \\ 1\; w.p.\, 1-\delta \end{cases}$
- Z_j is a Bernoulli random variable that represent the chance that the *jth* individual that did not dry out survives to become and adult w.p. λ.
- The number of surviving larvae that become adults from egg mass-*i* is W_i.
- Find the mean and variance of W_i.

Conditional mean

- Lower case variables are realizations of the random variable.
- Given, $Y = y$ and $X_i = x_i$, then the number of adults is $\sum_{j=1}^{yx_i} z_j = E\left[W_i | x_i, y, \sum z_j\right]$.
- $E\left[\sum_{j=1}^{yx_i} Z_j | x_i, y\right] = yx_i\lambda$, since $\sum Z_j \sim B(x_i y, \lambda(1-\lambda)x_i y)$
- $E[x_i Y\lambda | x_i] = x_i \cdot 0 \cdot \lambda \cdot \delta + x_i \cdot 1 \cdot \lambda \cdot (1-\delta) = x_i\lambda(1-\delta)$
- $E[W_i] = E[X_i\lambda(1-\delta)] = \mu_{X_i}\lambda(1-\delta)$

Conditional Variance

- Let $E[W_i] = \mu_{X_i}\lambda(1-\delta) = \widehat{W}$
- $Var(\sum_{j=1}^{yx_i} z_j | x_i, y) = E[yx_i\lambda(1-\lambda)] + Var(yx_i\lambda)$

$$= yx_i\lambda(1-\lambda) + E\{[yx_i\lambda - \widehat{w}]^2\}$$

$$E(Yx_i\lambda(1-\lambda)|x_i) + E\left[\left(Yx_i\lambda - \widehat{W}\right)^2 | x_i\right]$$

$$= x_i\,\lambda(1-\lambda)(1-\delta) + \widehat{W}^2\delta + (1-\delta)E\left[\left(x_i\lambda - \widehat{W}\right)^2\right]$$

$$= x_i\,\lambda(1-\lambda)(1-\delta) + \widehat{W}^2\delta + (1-\delta)E\left[\lambda^2 x_i^2 - 2\lambda x_i\widehat{W} + \widehat{W}^2\right]$$

$$\mu_x\lambda(1-\lambda)(1-\delta) + \widehat{W}^2\delta + (1-\delta)\left[\lambda^2\left(\mu_{x_i}^2 + \sigma_{x_i}^2\right) - 2\lambda\mu_{x_i}\widehat{W} + \widehat{W}^2\right]$$

$$= \cdots algebra = \widehat{W}\left(1 - \lambda - \widehat{W} + \lambda\, {}^{\sigma_{x_i}^2}\!/_{\mu_{x_i}} + \lambda\mu_{x_i}\right)$$

- Since this is for one clutch if the clutches are independent then we sum this quantity over all 1-4 clutches.
- The total variance could then range from 2.9 to 31. For the standard Wright-Fisher model this would be assumed to be 2.