Distributions	Density Function Form	Mean	Variance	Link Functions	Conj. Prior
Normal	$\frac{1}{\sqrt{2\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2	Identity	Normal/Gamma
Binomial	$\binom{n}{k}p^k(1-p)^{n-k}$	np	$np(1-p)$	Logit	Beta
Negative Binomial	$\binom{k+r-1}{k}(1-p)^r p^k$	$\frac{pr}{1-p}$	$\frac{pr}{1-p}$		Beta
Exponential	$\lambda e^{-\lambda x}$	λ^{-1}	λ^{-2}	Inverse	Gamma
Poisson	$\frac{\lambda^k e^{-\lambda}}{k!}$	λ	λ	Log	Gamma
Beta	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$		

General Notes

- The information matrix $\mathcal{I}$ is the variance-covariance matrix of the score statistics. It's inverse is the variance-covariance matrix for the β estimates.
- The score statistic U_j is the first derivative of the log-likelihood function with respect to β_j.
- Derivation of the linear regression coefficients using the method of maximum likelihood requires the normality assumptions.
- Derivation of the linear regression coefficients using method of least squares DOES NOT require the normality assumption.
- Deviances behave asymptotically according to a χ^2 distribution with degrees of freedom equal to the difference in degree of freedom between the log-likelihood functions involved.
- The expected value of the score statistic U is 0.
- Exponential Family Form: $\exp[a(y)b(\theta) + c(\theta) + d(y)]$
- The null deviance is 2 times the difference in likelihood of the saturated model and minimal model.
- The residual deviance is 2 times the difference in likelihood of the saturated model and them odel of interest.

Nominal Logistic Regression Model	Proportional Odds Model
$\log\frac{\pi_j}{\pi_1} = \beta_{0j} + \sum_{k=1}^{p-1}\beta_{kj}x_k$	$\log\frac{\sum_{k=1}^{j}\pi_k}{\sum_{k=j+1}^{K}\pi_k} = \beta_{0j} + \sum_{k=1}^{p-1}\beta_k x_k$

The deviance for a poisson model can be found with

$$D = 2\sum[o_i \log(o_i/e_i) - (o_i - e_i)]$$

where o_i is the observed data and e_i is the expected data.

Survival Analysis

- The survivor function $S(y) = 1 - F(y)$
- The hazard function $h(y) = f(y)/S(y) = -\frac{d}{dy}\log[S(y)]$
- The cumulative hazard rate function $H(y) = -\log[S(y)]$
- For an exponential, $S(y) = e^{-\theta y}, h(y) = \theta, H(y) = \theta y$
- For the Weibull, $h(y)$ is not constant. The exponential distribution is a special case of the weibull.
- Kaplan Meier estimate is the estimate for the survival function (i.e. $S(y)$).
- Parallel lines in log-cumulative hazard plots suggest the proportional hazards assumption may be appropriate where straight lines suggest weibull is a good choice.

Bayesian Inference : Example

θ	Hypothesis	$P(\theta)$ Prior	$P(y\|\theta)$ Likelihood	$P(y\|\theta)xP(\theta)$ Likelihood x Prior	$P(\theta\|y)$ Posterior
0.0	H_0	0.0333	0.0000	0.0000	0.0000
0.1	H_0	0.0333	0.0000	0.0000	0.0000
0.2	H_0	0.0333	0.0008	0.0000	0.0002
0.3	H_0	0.0333	0.0090	0.0003	0.0024
0.4	H_0	0.0333	0.0425	0.0014	0.0114
0.5	H_0	0.0333	0.1172	0.0039	0.0315
Sum		0.2000			0.0455
0.6	H_1	0.1600	0.2150	0.0344	0.2771
0.7	H_1	0.1600	0.2668	0.0427	0.3439
0.8	H_1	0.1600	0.2013	0.0322	0.2595
0.9	H_1	0.1600	0.0574	0.0092	0.0740
1.0	H_1	0.1600	0.0000	0.0000	0.0000
Sum		0.8000		.1241	0.9545

Prior in this case is 80% certain that $\theta > .5$ where likelihood $P(y|\theta)) \sim Bin(10, \theta)$. One can find the posterior probability

$$P(\theta|y) = \frac{P(y|\theta)P(\theta)}{\sum P(Y|\theta)P(\theta)}$$

Monte Carlo Integration, Markov Chains, MCMC

Metropolis Hastings Algorithm

- $\theta^* = \theta^{(i)} + Q$
- $\theta^{(i+1)} = \begin{cases} \theta^*, & \text{if } U < \alpha \\ \theta^{(i)}, & \text{otherwise} \end{cases}$
- $\alpha = \min\left\{\frac{P(\theta^*|y)}{P(\theta^{(i)}|y)}, 1\right\}$