C6: Discrete Distributions

	Poisson	Binomial	N-Binomial	Geometric
PMF	$p_k = e^{-\lambda}\dfrac{\lambda^k}{k!}$	$p_k = \begin{pmatrix} m \\ k \end{pmatrix} q^k (1-q)^{m-k}$	$p_k = \begin{pmatrix} k+r-1 \\ k \end{pmatrix}\left(\dfrac{\beta}{1+\beta}\right)^k\left(\dfrac{1}{1+\beta}\right)^r$ $k = 0,1,2\ldots, r>0, \beta>0$	$p_k = \left(\dfrac{\beta}{1+\beta}\right)^k\left(\dfrac{1}{1+\beta}\right)$ $\underline{k = 0,1,2\ldots, \beta>0}$
Moments	$E[X]=\lambda$ $Var(X)=\lambda$	$E[X]=mq$ $Var(X)=q(1-q)$	$E[X]=r\beta$ $Var(X)=r\beta(1+\beta)$	$E[X]=\beta$ $Var(X)=\beta(1+\beta)$
PMF/MGF	$G_Z = e^{\lambda(z-1)}$ $M_X = e^{\lambda(e^t-1)}$	$G_Z = (1-q+qz)^n$ $M_X = (1-q+qe^t)$	$G_Z = (1-\beta(z-1))^{-r}$ $M_X = (1-\beta(e^t-1))^{-r}$	$G_Z = (1-\beta(z-1))^{-1}$ $M_X = (1-\beta(e^t-1))^{-1}$

Likelihood Functions: $\lg \ell = \sum_{i=1}^{\infty} n_i \lg p(x_i \vert \theta)$		
	x_i	Observed datum
	n_i, $i = 1, 2\ldots k$	Number of times x_i is observed. $n_i = 1$ for point data.
	k	Number of groups

	Poisson	Binomial	N-Binomial	Geometric
Grouped Data *$lg\ l'(\theta)=0$*	$\dfrac{\partial \lg \ell}{\partial \lambda} = -n + \dfrac{n\bar{x}}{\lambda}$	$\dfrac{\partial \lg \ell}{\partial q} = \dfrac{n\bar{x}}{q} - \dfrac{mn - n\bar{x}}{1-q}$ $n = \sum_{i=0}^{k} n_i$	$\dfrac{\partial l}{\partial \beta} = \sum_{k=0}^{\infty} n_k\left(\dfrac{k}{\beta} + \dfrac{r-k}{1+\beta}\right)$ $\dfrac{\partial l}{\partial r} = -n_k \ln(1+\beta) + \sum_{k=i}^{\infty} n_k \sum_{m=0}^{k-1}\dfrac{1}{r+m}$	
MLEs	$\hat{\lambda} = \bar{X}$	$\hat{q} = \dfrac{1}{m}\dfrac{\sum_{k=0}^{\infty} k n_k}{\sum_{k=0}^{\infty} n_k}$	$\hat{\beta} = \dfrac{\bar{X}}{\hat{r}}$ r must be estimated numerically: $(Newton-Raphson)$ $r_k = r_{k-1} - \dfrac{H(r_{k-1})}{H'(r_{k-1})}$, $H(r_k) = n\lg\left(1+\dfrac{\bar{x}}{r}\right) - \sum_{k=1}^{\infty} n_k\left(\sum_{m=0}^{k-1}\dfrac{1}{r+m}\right) = 0$	$\hat{\beta} = \bar{X}$

	p_0	a	b	Parameter Space
Po (a, b, 0)	$e^{-\lambda}$			
ZT Po	0	0	λ	$\lambda > 0$
ZM Po	Arbitrary			
Bin	$(1-q)^m$			
ZT Bin	0	$-\dfrac{q}{1-q}$	$(m+1)\dfrac{q}{1-q}$	$0<q<1$ $m = 1,2\ldots$
ZM Bin	Arbitrary			
NB	$(1+\beta)^{-r}$			$r>0, \beta>0$
ZT NB	0	$\dfrac{\beta}{1+\beta}$	$(r-1)\dfrac{\beta}{1+\beta}$	$r>-1, r\neq 0, \beta>0$
ZM NB	Arbitrary			$r>-1, r\neq 0, \beta>0$
Geo	$(1+\beta)^{-1}$			
ZT Geo	0	$\dfrac{\beta}{1+\beta}$	0	$\beta > 0$
ZM Geo	Arbitrary			
Log	0	$\dfrac{\beta}{1+\beta}$	$-\dfrac{\beta}{1+\beta}$	$\beta > 0$
ZM Log	Arbitrary			

(a, b, 0) Class	(a, b, 1) Class
$p_k = \left(a + \dfrac{b}{k}\right)p_{k-1}, k = 1,2,3$ *Estimation*: $k\dfrac{\hat{p}_k}{\hat{p}_{k-1}} = k\dfrac{n_k}{n_{k-1}}$	$\dfrac{p_k}{p_{k-1}} = a + \dfrac{b}{k}, k = 2,3,4\ldots$ • **Zero-Truncated:** $p_0^T = 0$ (no observed data at zero) • **Zero-Modified:** $p_0^M > 0$ (Unusually large, or poor fit, number of observations at zero)

ZM Estimation	ZT Estimation
For some arbitrary constant c, $p_k^M = cp_k$ $G_Z^M = p_0^M + \sum_{k}^{\infty} p_k^M z^k$ $G_Z^M(1) = G_Z(1) = 1$ $\therefore c = \dfrac{1-p_0^M}{1-p_0}$	$Set: p_0^M = 0$ $P_z^T = \dfrac{P_z - p_0}{1-p_0}$ $p_z^T = \dfrac{p_k}{1-p_0}$ $p_z^M = (1-p_z^M)p_z^T$

Other on Poisson

Poisson as a limiting case of NB

$$\beta = \frac{\lambda}{r} \Rightarrow \lim_{r\to\infty}\left[1 - \frac{\lambda(z-1)}{r}\right]^{-r} = \ldots = e^{\lambda(z-1)}$$

$$\Rightarrow \lim_{r\to\infty} NB\left(r, \frac{r}{\lambda+r}\right) = Po(\lambda)$$