1.5 ESTUDO DE ALGUMAS DISTRIBUIÇÕES UNIVARIADAS

Há algumas famílias de variáveis aleatórias *discretas*, que pela sua importância prática têm nomes especiais. Iremos estudar alguns destes modelos nesta secção:

1.5.1 Distribuição binomial

Algumas experiências consistem numa sucessão de provas independentes podendo cada uma resultar num acontecimento ou no seu complementar. Por exemplo, um produto de uma determinada linha de montagem pode resultar em *defeituoso* ou *não defeituoso*, numa eleição cada voto pode ser *a favor* ou *contra* um candidato Sr. Silva, cada exame pode resultar numa *aprovação* ou numa *reprovação, et*c.

Estas experiências conhecidas como *provas de Bernoulli* apresentam as seguintes características:

1) A experiência consiste em **n** provas idênticas.

2) Cada prova resulta num acontecimento A (sucesso) ou no seu complementar A^c.

3) A probabilidade de sucesso em cada prova é igual a **p** e mantém-se constante de prova para prova. A probabilidade de insucesso é q=1-p.

4) As provas são independentes.

5) A v.a. associada a esta sucessão de provas é X - nº de sucessos observados nas n provas.

Definição: Uma v.a. X tem distribuição binomial de parâmetros n e p se a sua f.m.p. é igual a $P(X=k)=\binom{n}{k} p^k (1-p)^{n-k}$, $k=0,1,\ldots,n$ com $0 \le p \le 1$, e onde k é o nº de sucessos observados nas n provas, e designa-se por $X \cap Bi(n,p)$.

Exemplo 1: Suponhamos que numa população de 500 fusíveis 5% são defeituosos. Se uma amostra de 5 fusíveis é observada calcule a probabilidade de haver pelo menos um fusível defeituoso.

Seja A o acontecimento "o fusível é defeituoso", p= P(A)= 0.05, X-nº de fusíveis defeituosos na amostra, é uma v.a. com distribuição Bi(5, 0.05). Então,

$$P(X \ge 1) = 1 - P(X=0) = 1 - \binom{5}{0}(0.05)^0 (0.95)^5 = 1 - 0.774 = 0.226.$$

Esta variável tem valor médio e variância finitos.

$$E(X)=\sum_{k=0}^{n} kP(X=k)=\sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k}=\sum_{k=0}^{n} k\frac{n!}{(n-k)!\,k!}p^k(1-p)^{n-k}=$$

$$=\sum_{k=1}^{n}\frac{n(n-1)!}{(n-k)!(k-1)!}pp^{k-1}(1-p)^{n-k}=np\sum_{k=1}^{n-1}\binom{n-1}{k-1}p^{k-1}(1-p)^{n-k}=$$

$$=np(p+1-p)^{n-1}=np$$

de igual modo podíamos calcular $E(X^2)= np(1-p)+(np)^2$ e consequentemente teremos $Var(X)=np(1-p)$.

Isto é, se $X\cap Bi(n,p)$ então $\mathrm{E}(X)=np$ e $\mathrm{Var}(X)=np(1-p)$, o que mostra que os parâmetros da distribuição estão relacionados com as suas medidas de localização e dispersão.

NOTA: No caso do exemplo anterior o nº médio de fusíveis defeituosos numa amostra de 5 fusíveis é 5x0.05=0.25 e a variância é 5x0.05x0.95=0.2375.
Como acontece na maioria dos casos o valor médio NÃO é um valor assumido pela v.a. X-nº de fusíveis defeituosos numa amostra de 5.

Exemplo 2: Suponhamos agora que se recolheu uma amostra de 50 fusíveis e queremos calcular a probabilidade de haver pelo menos 5 fusíveis defeituosos nessa amostra.
Teremos como anteriormente X- nº de fusíveis defeituosos numa amostra de 50, onde agora $X\cap Bi(50,0.05)$ e a probabilidade pedida é igual a

$$P(X\geq 5)=1-P(X<5)=1-\sum_{i=0}^{4}P(X=i)=1-\sum_{i=0}^{4}\binom{50}{i}(0.05)^i(0.95)^{50-i}$$

Dado que a dimensão da amostra é grande, torna-se fastidioso calcular as probabilidades que constituem as parcelas do somatório.
A distribuição que se segue ajuda a resolver o problema nestas circunstâncias.

1.5.2 Distribuição de Poisson

Considerando o modelo probabilístico anterior é fácil mostrar que podemos obter um valor aproximado de $p_k= P(X= k)$ quando n é "grande" e p toma valores muito pequenos.

$$
\begin{aligned}
P(X=k) &= \frac{n!}{(n-k)!\,k!}\, p^k (1-p)^{n-k} = \\
&= \frac{1}{(1-p)^k k!}\, p^k (1-p)^n n(n-1)(n-2)\ldots(n-k+1) = \\
&= \frac{1}{(1-p)^k k!}\, p^k (1-p)^n n^k \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{k-1}{n}\right) \approx \\
&\approx \frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n \underset{n}{\rightarrow} e^{-\lambda}\frac{\lambda^k}{k!}
\end{aligned}
$$

onde $np \approx \lambda$, quando $n \to +\infty$ e $p \approx 0$

Através do resultado anterior podemos facilmente calcular a probabilidade pedida no exemplo 2, onde $\lambda \approx 50 \times 0.05 = 2.5$. Assim,

$$P(X \geq 5) \approx 1 - \sum_{i=0}^{4} e^{-2.5}\frac{2.5^i}{i!} = 0.10882$$

A colecção de valores

$p_k = e^{-\lambda}\dfrac{\lambda^k}{k!}$, $k = 0,1,2,\ldots$, $p_k \geq 0$ e $\displaystyle\sum_{k=0}^{+\infty} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = 1$ constitui uma f.m.p. de uma v.a. .

Definição: Uma variável aleatória X que toma valores em $\mathbb{N}_0$, com f.m.p. dada por $P(X=k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$, $k = 0,1,2,\ldots$, tem distribuição de Poisson de parâmetro $\lambda > 0$ e designa-se por $X \cap P(\lambda)$.

O parâmetro λ desta distribuição tem um significado especial como podemos ver. Assim, calcule-se

$$E(X) = \sum_{i=0}^{+\infty} x_i p_i = \sum_{i=0}^{+\infty} i p_i = \sum_{i=0}^{+\infty} i e^{-\lambda}\frac{\lambda^i}{i!} = \lambda e^{-\lambda}\sum_{i=1}^{+\infty}\frac{\lambda^{i-1}}{(i-1)!} = \lambda e^{-\lambda}\sum_{j=0}^{+\infty}\frac{\lambda^j}{j!} = \lambda e^{-\lambda}e^{\lambda} = \lambda$$

de forma análoga se calcula

$$Var(X) = E\left(X^2\right) - E^2(X) = \lambda^2 + \lambda - \lambda^2$$
$$Var(X) = \sigma^2(X) = \lambda$$