Distribuzione Binomiale

Introduzione

Consideriamo un processo (o prova) che può portare a:

2 eventi mutuamente esclusivi

es: Nascita → Maschio ♂

es: Nascita → Femmina ♀

Hp:

1. Le probabilità di un tipo di evento (e quindi anche dell'altro …) sia COSTANTE (cioè di prova in prova non cambi)
2. Le prove siano INDIPENDENTI (cioè l'esito di una non influenzi quello di un'altra)
3. Gli eventi (risultato della prova) siano MUTUAMENTE ESCLUSIVI

Questo tipo di processo si chiama BERNOULLIANO

↓

La distribuzione di probabilità ad esso legata si chiama BINOMIALE

Esempio

In una certa popolazione (ad es. Italiana) la probabilità di nascere Maschio è ≈ 0.51 ed essa è costante (negli ultimi anni)

Qual è la probabilità che su 3 figli 2 siano maschi?

Possibili sequenze

M M F

M F M

F M M

Su $2^3=8$ possibili sequenze

Ciascuna di queste sequenze ha la stessa probabilità di verificarsi che è:

$$0.51 \cdot 0.51 \cdot 0.49$$
$$P(M) \times P(M) \times P(F)$$

Quindi la probabilità di ottenere MMF o MFM o FMM è la somma della probabilità di ciascuna sequenza

$$3 \cdot \left[(0.51)^2 \cdot 0.49\right]$$

COEFFICIENTE BINOMIALE

Come possiamo contare tutte le possibili combinazione senza doverle definire tutte?

Esiste una formula matematica che permette di contare le possibili **combinazioni** di **x** oggetti su un insieme di **n** :

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

<u>Coefficiente binomiale</u>: denota il numero di sequenze di lunghezza n che si possono formare con x elementi aventi il carattere d ed (n-x) elementi aventi il carattere đ.

Dove: n! si legge "n fattoriale" ed
n!=n*(n-1)*(n-2)*(n-3)*…*1
0!=1

Per esempio le possibili combinazioni di 2 (=x) maschi in 3 (=n) nascite sono:

$$\binom{n}{x} = \frac{3!}{2!\,(3-2)!} = \frac{3*2*1}{2*1*1} = 3$$

Due tipi di eventi → a
Due tipi di eventi → b

n prove

Ci sono 2^n possibili sequenze distinte, ciascuna può avere
x elementi di tipo a

Con n=3 ho 8 possibili sequenze:

1	3 "a"	x=3	aaa
3	2 "a"	x=2	aab aba baa
3	1 "a"	x=1	abb bab bba
1	0 "a"	x=0	bbb

Con n=2 ho 4 possibili sequenze:

1	2 "a"	x=2	aa
2	1 "a"	x=1	ab ba
1	0 "a"	x=0	bb

Con n=1 ho 2 possibili sequenze:

1	1 "a"	x=1	a
1	0 "a"	x=0	b

Triangolo di Tartaglia (1)

$(a+b)^1$	1 1
$(a+b)^2$	1 2 1
$(a+b)^3$	1 3 3 1
$(a+b)^4$	1 4 6 4 1

$$(a+b)^2 = (a+b)\cdot(a+b) = (aa+ab+ba+bb) = a^2+b^2+2ab$$

$$(a+b)^3 = a^3+b^3+3a^2b+3ab^2$$

Triangolo di Tartaglia (2)

$$
\begin{array}{c|cccccccc}
 & x & 0 & 1 & 2 & 3 & 4 & \cdots & x \\
\hline
n & & & & & & & & \\
0 & & 1 & & & & & & \\
1 & & 1 & 1 & & & & & \\
2 & & 1 & 2 & 1 & & & & \\
3 & & 1 & 3 & 3 & 1 & & & \\
4 & & 1 & 4 & 6 & 4 & 1 & & \\
\vdots & & \vdots & & & & & \ddots & \\
n & & & \cdots & & \cdots & & \cdots & \binom{n}{x}
\end{array}
$$

Distribuzione binomiale

Si abbia un processo per la produzione di compresse tale che la probabilità di una compressa di essere difettosa sia $\pi=0.01$.

Si vuole calcolare la probabilità di trovare 0,1,2,3,4 compresse difettose in un blister di 4 compresse, sotto l'assunto che la probabilità di una compressa di finire in un dato blister non dipenda dalle caratteristiche della compressa stessa.

Si noti che il blister può essere considerato un campione tratto da un universo (la produzione di compresse) virtualmente infinito.

… continua

Per enumerazione dei casi possibili, si osserva che 4 compresse difettose (d) e no (đ) si combinano, con probabilità π ed (1-π), in 2^4=16 diversi modi.

In analogia con ciò che si è già visto, la probabilità di osservare un blister con x compresse difettose è la somma delle probabilità associate ad ognuno dei diversi modi con cui il blister può includere x compresse difettose. Ad esempio, vi sono 4 modi con cui un blister può avere una sola compressa difettosa: essa può essere la prima, la seconda, la terza o la quarta.

Distribuzione binomiale (2)

- Variabile x={numero di compresse difettose in un blister da 4 compresse}

- π=0.01 è la probabilità del processo di generare una compressa difettosa.

	compresse			
x	1	2	3	4
0	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\bar{d}$
1	d	$\bar{d}$	$\bar{d}$	$\bar{d}$
	$\bar{d}$	d	$\bar{d}$	$\bar{d}$
	$\bar{d}$	$\bar{d}$	d	$\bar{d}$
	$\bar{d}$	$\bar{d}$	$\bar{d}$	d
2	d	d	$\bar{d}$	$\bar{d}$
	d	$\bar{d}$	d	$\bar{d}$
	d	$\bar{d}$	$\bar{d}$	d
	$\bar{d}$	d	d	$\bar{d}$
	$\bar{d}$	d	$\bar{d}$	d
	$\bar{d}$	$\bar{d}$	d	d
3	d	d	d	$\bar{d}$
	d	d	$\bar{d}$	d
	d	$\bar{d}$	d	d
	$\bar{d}$	d	d	d
4	d	d	d	d

Distribuzione binomiale (2)

- Variabile x={numero di compresse difettose in un blister da 4 compresse}
- π=0.01 è la probabilità del processo di generare una compressa difettosa.

	compresse				
x	1	2	3	4	PROBABILITÀ
0	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\bar{d}$	$(1-\pi)(1-\pi)(1-\pi)(1-\pi)$
1	d	$\bar{d}$	$\bar{d}$	$\bar{d}$	$\pi\ (1-\pi)(1-\pi)(1-\pi)$
	$\bar{d}$	d	$\bar{d}$	$\bar{d}$	$(1-\pi)\ \pi\ (1-\pi)(1-\pi)$
	$\bar{d}$	$\bar{d}$	d	$\bar{d}$	$(1-\pi)(1-\pi)\ \pi\ (1-\pi)$
	$\bar{d}$	$\bar{d}$	$\bar{d}$	d	$(1-\pi)(1-\pi)(1-\pi)\ \pi$
2	d	d	$\bar{d}$	$\bar{d}$	$\pi\ \pi\ (1-\pi)(1-\pi)$
	d	$\bar{d}$	d	$\bar{d}$	$\pi\ (1-\pi)\ \pi\ (1-\pi)$
	d	$\bar{d}$	$\bar{d}$	d	$\pi\ (1-\pi)(1-\pi)\ \pi$
	$\bar{d}$	d	d	$\bar{d}$	$(1-\pi)\ \pi\ \pi\ (1-\pi)$
	$\bar{d}$	d	$\bar{d}$	d	$(1-\pi)\ \pi\ (1-\pi)\ \pi$
	$\bar{d}$	$\bar{d}$	d	d	$(1-\pi)(1-\pi)\ \pi\ \pi$
3	d	d	d	$\bar{d}$	$\pi\ \pi\ \pi\ (1-\pi)$
	d	d	$\bar{d}$	d	$\pi\ \pi\ (1-\pi)\ \pi$
	d	$\bar{d}$	d	d	$\pi\ (1-\pi)\ \pi\ \pi$
	$\bar{d}$	d	d	d	$(1-\pi)\ \pi\ \pi\ \pi$
4	d	d	d	d	$\pi\ \pi\ \pi\ \pi$

Distribuzione binomiale (2)

- Variabile x={numero di compresse difettose in un blister da 4 compresse}
- π=0.01 è la probabilità del processo di generare una compressa difettosa.

x	compresse 1 2 3 4	PROBABILITÀ			
0	$\bar{d}\ \bar{d}\ \bar{d}\ \bar{d}$	$(1-\pi)(1-\pi)(1-\pi)(1-\pi)$	$(1-\pi)^4$ =	$.99^4$ =	.96059601
1	$d\ \bar{d}\ \bar{d}\ \bar{d}$	$\pi\ (1-\pi)(1-\pi)(1-\pi)$ =	$\pi(1-\pi)^3$ =	$.01^1\times.99^3$ =	.00970299
	$\bar{d}\ d\ \bar{d}\ \bar{d}$	$(1-\pi)\ \pi\ (1-\pi)(1-\pi)$ =	$\pi(1-\pi)^3$ =	$.01^1\times.99^3$ =	.00970299
	$\bar{d}\ \bar{d}\ d\ \bar{d}$	$(1-\pi)(1-\pi)\ \pi\ (1-\pi)$ =	$\pi(1-\pi)^3$ =	$.01^1\times.99^3$ =	.00970299
	$\bar{d}\ \bar{d}\ \bar{d}\ d$	$(1-\pi)(1-\pi)(1-\pi)\ \pi$ =	$\pi(1-\pi)^3$ =	$.01^1\times.99^3$ =	.00970299
		=	$4\pi(1-\pi)^3$ =	$4\times.01^1\times.99^3$ =	.03881196
2	$d\ d\ \bar{d}\ \bar{d}$	$\pi\ \pi\ (1-\pi)(1-\pi)$ =	$\pi^2(1-\pi)^2$ =	$.01^2\times.99^2$ =	.00009801
	$d\ \bar{d}\ d\ \bar{d}$	$\pi\ (1-\pi)\ \pi\ (1-\pi)$ =	$\pi^2(1-\pi)^2$ =	$.01^2\times.99^2$ =	.00009801
	$d\ \bar{d}\ \bar{d}\ d$	$\pi\ (1-\pi)(1-\pi)\ \pi$ =	$\pi^2(1-\pi)^2$ =	$.01^2\times.99^2$ =	.00009801
	$\bar{d}\ d\ d\ \bar{d}$	$(1-\pi)\ \pi\ \pi\ (1-\pi)$ =	$\pi^2(1-\pi)^2$ =	$.01^2\times.99^2$ =	.00009801
	$\bar{d}\ d\ \bar{d}\ d$	$(1-\pi)\ \pi\ (1-\pi)\ \pi$ =	$\pi^2(1-\pi)^2$ =	$.01^2\times.99^2$ =	.00009801
	$\bar{d}\ \bar{d}\ d\ d$	$(1-\pi)(1-\pi)\ \pi\ \pi$ =	$\pi^2(1-\pi)^2$ =	$.01^2\times.99^2$ =	.00009801
			$6\pi^2(1-\pi)^2$	$6\times.01^2\times.99^2$	.00058806
3	$d\ d\ d\ \bar{d}$	$\pi\ \pi\ \pi\ (1-\pi)$ =	$\pi^3(1-\pi)$ =	$.01^3\times.99^1$ =	.00000099
	$d\ d\ \bar{d}\ d$	$\pi\ \pi\ (1-\pi)\ \pi$ =	$\pi^3(1-\pi)$ =	$.01^3\times.99^1$ =	.00000099
	$d\ \bar{d}\ d\ d$	$\pi\ (1-\pi)\ \pi\ \pi$ =	$\pi^3(1-\pi)$ =	$.01^3\times.99^1$ =	.00000099
	$\bar{d}\ d\ d\ d$	$(1-\pi)\ \pi\ \pi\ \pi$ =	$\pi^3(1-\pi)$ =	$.01^3\times.99^1$ =	.00000099
			$4\pi^3(1-\pi)$	$4\times.01^3\times.99^1$	.00000396
4	$d\ d\ d\ d$	$\pi\ \pi\ \pi\ \pi$ =	π^4 =	$.01^4$ =	.00000001

Distribuzione binomiale (3)

Detto n il numero delle occasioni (ad es. il numero di compresse nel blister) ed x il numero di eventi (ad es. il numero di compresse difettose nel blister) ognuno dei quali ha la stessa probabilità di verificarsi, si ha che:

$$P(x|n) = \frac{n!}{x! \cdot (n-x)!} \cdot \pi^x \cdot (1-\pi)^{n-x} =$$

$$= \binom{n}{x} \cdot \pi^x \cdot (1-\pi)^{n-x}$$

La variabile le cui probabilità sono assegnate secondo tale legge è chiamata variabile casuale binomiale e si denota come

$x \sim Bi(\pi, n)$.

… continua

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

<u>Coefficiente binomiale</u>: denota il numero di sequenze di lunghezza n che si possono formare con x elementi aventi il carattere d ed (n-x) elementi aventi il carattere đ.

$$\pi^{x} \cdot (1-\pi)^{n-x}$$

Rappresenta la probabilità che si verifichino <u>contemporaneamente</u> x volte l'evento con probabilità π e (n-x) volte l'evento con probabilità (1- π).

Se è nota la funzione di probabilità, possiamo assegnare le probabilità agli eventi senza dover enumerare i casi possibili.

Nell'esempio la probabilità di osservare un blister con 2 compresse difettose può essere calcolata direttamente:

$$P(2|4) = \binom{4}{2} \cdot 0.01^{2} \cdot (1-0.01)^{4-2} = \frac{4!}{2! \times (4-2)!} \cdot 0.01^{2} \cdot 0.99^{2} =$$

$$= \frac{24}{4} \cdot 0.0001 \cdot 0.9801 = 0.00058806$$

… riepilogando …

La distribuzione binomiale mostra la probabilità di diversi risultati per una serie di eventi casuali, ognuno dei quali può assumere solo uno tra due valori

Esempio di distribuzione binomiale (1)

Ad un centro trasfusionale afferiscono ogni mattina 10 pazienti talassemici per ricevere una sacca di sangue. Se vi è una probabilità del 10% che un paziente sviluppi una reazione febbrile (RF), qual è la probabilità di osservare in una mattina 0,1,…,5,…,10 pazienti che sviluppano RF?

Def.

x={numero di pazienti che sviluppano reazione febbrile in seguito a trasfusione di una sacca di sangue su 10 pazienti trasfusi}

n=10 e π=0.10

$$P(x|10) = \binom{10}{x} \cdot 0.10^{x} \cdot (1-0.10)^{10-x}$$

x			
0			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

… continua

$x = 0$

$$P(0|10) = \binom{10}{0} \cdot 0.10^0 \cdot (1-0.10)^{10-0} = \frac{10!}{\underbrace{0!}_{1} \cdot 10!} \cdot \underbrace{0.10^0}_{1} \cdot (0.90)^{10} = (0.90)^{10} = 0.3487$$

$\boxed{0! = 1 \quad \text{Per convenzione}}$

$x = 1$

$$P(1|10) = \binom{10}{1} \cdot 0.10^1 \cdot (1-0.10)^{10-1} = \frac{10!}{1! \cdot 9!} \cdot 0.10^1 \cdot (0.90)^9 =$$

$$= \frac{10 \cdot 9!}{9!} 0.10 \cdot (0.90)^9 = 10 \cdot 0.10 \cdot (0.90)^9 = 0.3874$$

$x = 5$

$$P(5|10) = \binom{10}{5} \cdot 0.10^5 \cdot (1-0.10)^{10-5} = \frac{10!}{5! \cdot 5!} \cdot 0.10^5 \cdot (0.90)^5 =$$

$$= \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5!}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 \cdot 5!} 0.10^5 \cdot (0.90)^5 = 252 \cdot 0.10^5 \cdot (0.90)^5 = 0.0015$$

… continua

Ad un centro trasfusionale afferiscono ogni mattina 10 pazienti talassemici per ricevere una sacca di sangue. Se vi è una probabilità del 10% che un paziente sviluppi una reazione febbrile (RF), qual è la probabilità di osservare in una mattina 0,1,…,5,…,10 pazienti che sviluppano RF?

Def.

x={numero di pazienti che sviluppano reazione febbrile in seguito a trasfusione di una sacca di sangue su 10 pazienti trasfusi}

n=10 e π=0.10

$$P(x\,|\,10) = \binom{10}{x} \cdot 0.10^{x} \cdot (1 - 0.10)^{10-x}$$

x	P(X=x)		
0	0.34868		
1	0.38742		
2	0.19371		
3	0.05740		
4	0.01116		
5	0.00149		
6	0.00013		
7	0.00001		
8	0.00000		
9	0.00000		
10	0.00000		

Esercizio: calcolare

1. P{almeno 1 paziente sviluppi RF}
2. P{non più di 1 paziente sviluppi RF}
3. P{più di 3 pazienti sviluppino RF}

$$1.\ P\{X \geq 1\} = P\{X = 1 \cup X = 2 \cup \ldots \cup X = 10\} = \sum_{i=1}^{10} P\{X = i\} = 1 - P\{X = 0\} = 1 - 0.349 = 0.651$$

$$2.\ P\{X \leq 1\} = P\{X = 0 \cup X = 1\} = P\{X = 0\} + P\{X = 1\} = 0.349 + 0.387 = 0.736$$

$$3.\ P\{X \geq 4\} = P\{X = 4 \cup X = 5 \cup \ldots \cup X = 10\} = \sum_{i=4}^{10} P\{X = i\} = 1 - \sum_{j=0}^{3} P\{X = j\} =$$

$$= 1 - 0.349 - 0.387 - 0.194 - 0.057 = 0.013$$

… continua

Ad un centro trasfusionale afferiscono ogni mattina 10 pazienti talassemici per ricevere una sacca di sangue. Se vi è una probabilità del 10% che un paziente sviluppi una reazione febbrile (RF), qual è la probabilità di osservare in una mattina 0,1,…,5,…,10 pazienti che sviluppano RF?

<u>Def.</u>

x={numero di pazienti che sviluppano reazione febbrile in seguito a trasfusione di una sacca di sangue su 10 pazienti trasfusi}

n=10 e π=0.10

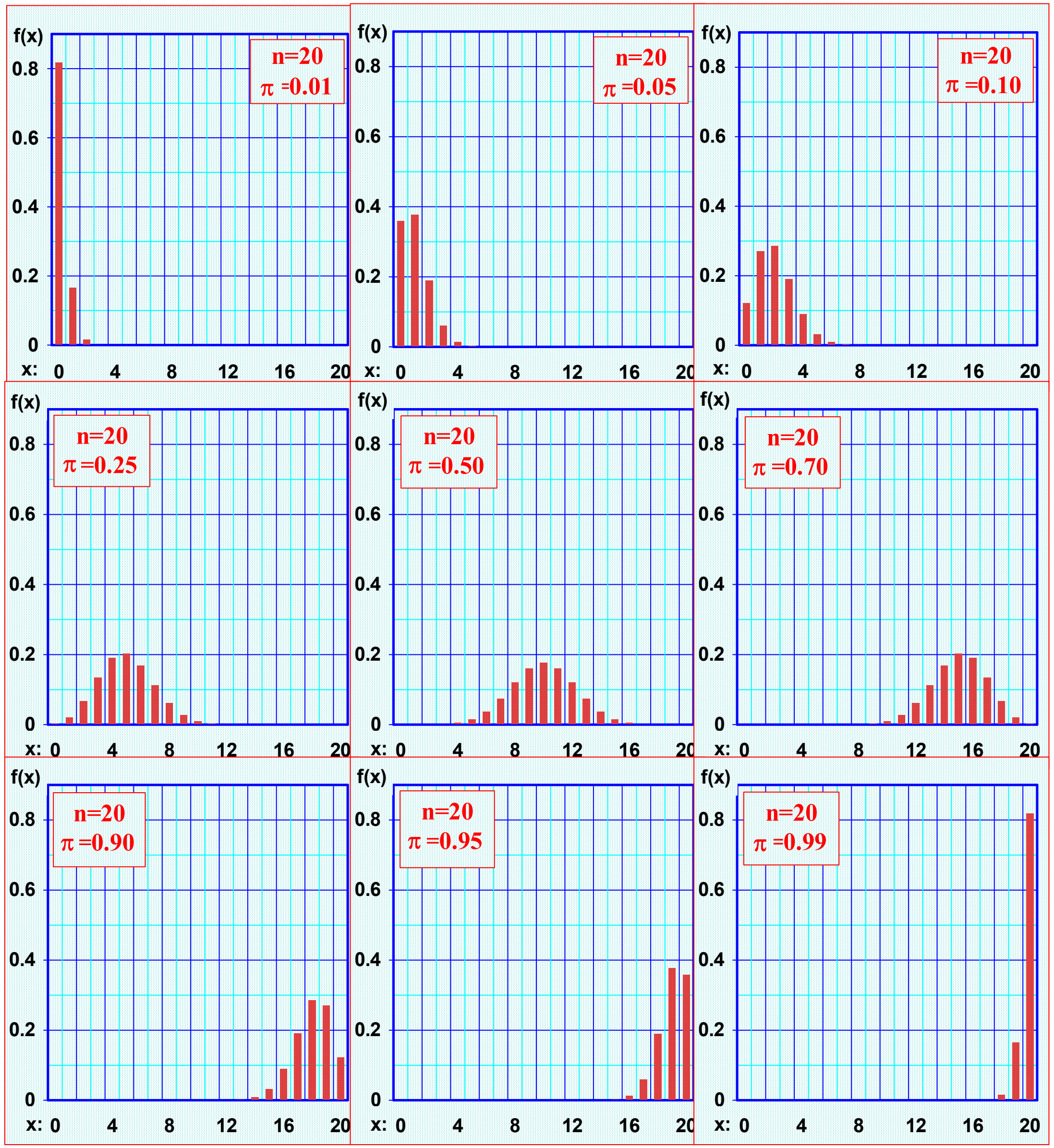
$$P(x\,|\,10) = \binom{10}{x} \cdot 0.10^{x} \cdot (1-0.10)^{10-x}$$

x	P(X=x)	P(X≤x)	1-P(X≤x)
0	0.34868	0.34868	1.00000
1	0.38742	0.73610	0.65132
2	0.19371	0.92981	0.26390
3	0.05740	0.98721	0.07019
4	0.01116	0.99837	0.01279
5	0.00149	0.99986	0.00163
6	0.00013	0.99999	0.00014
7	0.00001	1.00000	0.00001
8	0.00000	1.00000	0.00000
9	0.00000	1.00000	0.00000
10	0.00000	1.00000	0.00000

Forma della distribuzione binomiale

Variabili casuali binomiali, tutte con parametro n=20, ma con differente parametro π

(π=0.01, 0.05, … , 0.99).



La distribuzione è:

simmetrica

π=0.50

asimmetrica positiva

$\pi \to 0$

asimmetrica negativa

$\pi \to 1$

DISTRIBUZIONE BINOMIALE
$X \sim Bi(\pi=0.1, n=10)$

- Quanti pazienti con reazioni febbrili mi aspetto in un giorno?

$$E(X)= \pi * n= 10*0.1=1$$

- Che variabilità ha il numero di pazienti con RF in un giorno?

$$VAR(X)=n*\pi*(1- \pi)=10*0.1*0.9 = 0.9$$

Esercizio

Un prodotto farmaceutico provoca un grave effetto collaterale a 3 pazienti su 100. Un'industria farmaceutica desidera sottoporre a prova il medicinale.

1) Qual è la probabilità che l'effetto collaterale si verifichi al massimo in 1 soggetto in un campione casuale di 20 pazienti che prendono il medicinale?
2) Qual è la probabilità che tutti i 20 soggetti manifestino l’effetto collaterale?
3) Quanti pazienti con l’effetto collaterale mi aspetto sui 20 soggetti?

Risposte

1) P(X≤1)=P(X=1)+P(X=0)

$$P(X=0)=\binom{20}{0}0{,}03^{0}0{,}97^{20}=0{,}5438$$

$$P(X=1)=\binom{20}{1}0{,}03^{1}0{,}97^{19}=\frac{20*19!}{19!\,1!}0{,}03^{1}0{,}97^{19}=20*0{,}03^{1}0{,}97^{19}$$
$$=0{,}3364$$

$$P(X\le 1)=P(X=0)+P(X=1)=0{,}5438+0{,}3364=0{,}8802$$

2) $P(X=20)=\binom{20}{20}0{,}03^{20}0{,}97^{0}=0{,}03^{20}=3.486784*10^{-31}\approx 0$

3) E(X)=n*π=20*0,03=0,6