LE PRINCIPALI

DISTRIBUZIONI DI PROBABILITÀ

SOMMARIO

- **Definizione di distribuzione di probabilità**
- **Le principali distribuzioni discrete**
- **Le principali distribuzioni continue**
- **L'approssimazione normale di alcune distribuzioni discrete**

DEFINIZIONE DI DISTRIBUZIONE DI PROBABILITÀ

Una **distribuzione di probabilità** è un modello matematico che collega i valori di una variabile alle probabilità che tali valori possano essere osservati.

Le distribuzioni di probabilità vengono utilizzate per modellizzare il comportamento di un fenomeno di interesse in relazione alla popolazione di riferimento, ovvero alla totalità dei casi di cui lo sperimentatore osserva un dato campione.

In questo contesto la variabile di interesse è vista come una **variabile casuale** (o **variabile aleatoria, v.a.**) la cui legge di probabilità esprime il grado di incertezza con cui i suoi valori possono essere osservati.

DISTRIBUZIONI DI PROBABILITÀ DISCRETE E CONTINUE

In base alla scala di misura della variabile di interesse *X*, possiamo distinguere due tipi di distribuzioni di probabilità:

1. **distribuzioni continue**: la variabile viene espressa su un scala continua (es: il diametro del pistone)
2. **distribuzioni discrete**: la variabile viene misurata con valori numerici interi (es: numero di elementi non conformi o difettosi in un circuito stampato)

Formalmente, le distribuzioni di probabilità vengono espresse da una legge matematica detta **funzione di densità di probabilità** (indicata con $f(x)$) o **funzione di probabilità** (indicata con $p(x)$) rispettivamente per le distruzioni continue o discrete.

DISTRIBUZIONI DI PROBABILITÀ DISCRETE E CONTINUE

Esempio di **funzione di probabilità**

Esempio di **funzione di densità di probabilità**

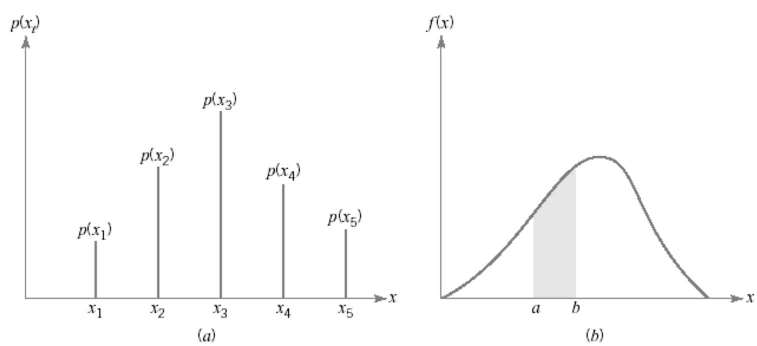


Figure 2-9 Probability distributions. (*a*) Discrete case. (*b*) Continuous case.

ESEMPIO DI DISTRIBUZIONE DI PROBABILITÀ DISCRETA

EXAMPLE 2-5

A Discrete Distribution

A manufacturing process produces thousands of semiconductor chips per day. On the average, 1% of these chips do not conform to specifications. Every hour, an inspector selects a random sample of 25 chips and classifies each chip in the sample as conforming or nonconforming. If we let x be the random variable representing the number of nonconforming chips in the sample, then the probability distribution of x is

$$p(x) = \binom{25}{x}(0.01)^x(0.99)^{25-x} \qquad x = 0, 1, 2, \ldots, 25$$

where $\binom{25}{x} = 25![x!(25-x)!]$. This is a *discrete* distribution, since the observed number of nonconformances is $x = 0, 1, 2, \ldots, 25$, and is called the **binomial distribution.** We may calculate the probability of finding one or fewer nonconforming parts in the sample as

$$\begin{aligned} P(x \le 1) &= P(x=0) + P(x=1) \\ &= p(0) + p(1) \\ &= \sum_{x=0}^{1} \binom{25}{x}(0.01)^x(0.99)^{25-x} \\ &= \frac{25!}{0!\,25!}(0.99)^{25}(0.01)^0 + \frac{25!}{1!\,24!}(0.99)^{24}(0.01)^1 \\ &= 0.7778 + 0.1964 = 0.9742 \end{aligned}$$

ESEMPIO DI DISTRIBUZIONE DI PROBABILITÀ CONTINUA

DISTRIBUZIONE UNIFORME

EXAMPLE 2-6

A Continuous Distribution

Suppose that x is a random variable that represents the actual contents in ounces of a 1-1b bag of coffee beans. The probability distribution of x is assumed to be

$$f(x) = \frac{1}{1.5} \qquad 15.5 \le x \le 17.0$$

This is a *continuous* distribution, since the range of x is the interval [15.5, 17.0]. This distribution is called the **uniform distribution,** and it is shown graphically in Fig. 2-10. Note that the area under the function $f(x)$ corresponds to probability, so that the probability of a bag containing less than 16.0 oz is

$$P\{x \le 16.0\} = \int_{15.5}^{16.0} f(x)dx = \int_{15.5}^{16.0} \frac{1}{1.5}dx = \left.\frac{x}{1.5}\right|_{15.5}^{16.0} = \frac{16.0-15.5}{1.5} = 0.3333$$

This follows intuitively from inspection of Fig. 2-9.

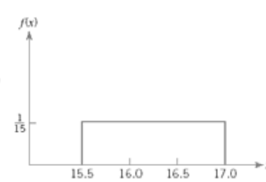


Figure 2-10 The uniform distribution for Example 2-6.

MEDIA E VARIANZA DI UNA DISTRIBUZIONE DI PROBABILITÀ

La media (o valore atteso) μ e la varianza σ^2 (deviazione standard σ) di una v.a. *X* sono i parametri di maggiore interesse della distribuzione di probabilità di *X*, in quanto essi esprimono rispettivamente la tendenza centrale e la variabilità della v.a. *X*.

$$\mu = \begin{cases} \int_{-\infty}^{\infty} x f(x)\, dx, x \text{ continuous} \\ \sum_{i=1}^{\infty} x_i p(x_i), x \text{ discrete} \end{cases}$$

$$\sigma^2 = \begin{cases} \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx, x \text{ continuous} \\ \sum_{i=1}^{\infty} (x_i-\mu)^2 p(x_i), x \text{ discrete} \end{cases}$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}}$$

DISTRIBUZIONE BINOMIALE

È il modello adatto per il campionamento da una popolazione infinita, dove *p* rappresenta la frazione di elementi difettosi o non conformi presenti nella popolazione.

Definition

The **binomial distribution** with parameters $n \geq 0$ and $0 < p < 1$ is

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0,1,\ldots,n \qquad (2\text{-}11)$$

The mean and variance of the binomial distribution are

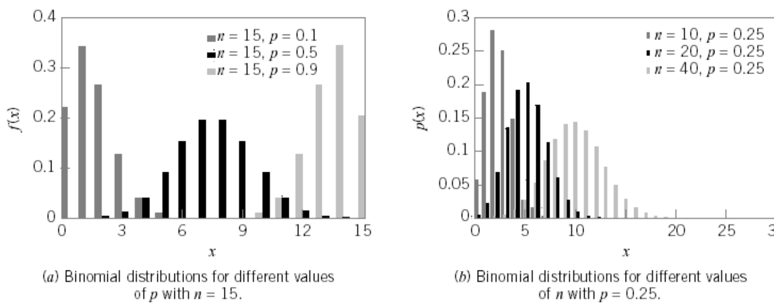
$$\mu = np \qquad (2\text{-}12)$$

and

$$\sigma^2 = np(1-p) \qquad (2\text{-}13)$$

DISTRIBUZIONE BINOMIALE

La variabile casuale X rappresenta il numero di "successi" in n prove indipendenti di Bernoulli con probabilità di successo costante pari a *p* in ogni prova.



(*a*) Binomial distributions for different values of p with $n = 15$.

(*b*) Binomial distributions for different values of n with $p = 0.25$.

Figure 2-14 Binomial distributions for selected values of n and p.

DISTRIBUZIONE BINOMIALE

A random variable that arises frequently in statistical quality control is

$$\hat{p} = \frac{x}{n} \tag{2-14}$$

where x has a binomial distribution with parameters n and p. Often $\hat{p}$ is the ratio of the observed number of defective or nonconforming items in a sample (x) to the sample size (n) and this is usually called the **sample fraction defective** or **sample fraction nonconforming.** The "ˆ" symbol is used to indicate that $\hat{p}$ is an estimate of the true, unknown value of the binomial parameter p. The probability distribution of $\hat{p}$ is obtained from the binomial, since

$$P\{\hat{p} \le a\} = P\left\{\frac{x}{n} \le a\right\} = p\{x \le na\} = \sum_{x=0}^{[na]} \binom{n}{x} p^x (1-p)^{n-x}$$

where $[na]$ denotes the largest integer less than or equal to na. It is easy to show that the mean of $\hat{p}$ is p and that the variance of $\hat{p}$ is

$$\sigma_{\hat{p}}^2 = \frac{p(1-p)}{n}$$

DISTRIBUZIONE POISSON

È il modello adatto per modellare il numero di difetti o non conformità che si trovano in una unità di prodotto.

Definition

The **Poisson distribution** is

$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0, 1, \ldots \qquad (2\text{-}15)$$

where the parameter $\lambda > 0$. The **mean** and **variance** of the Poisson distribution are

$$\mu = \lambda \qquad (2\text{-}16)$$

and

$$\sigma^2 = \lambda \qquad (2\text{-}17)$$

DISTRIBUZIONE POISSON

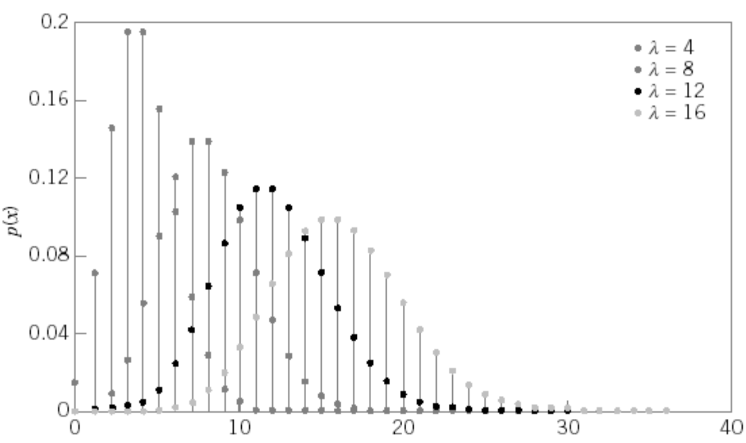


Figure 2-15 Poisson probability distributions for selected values of λ.

DISTRIBUZIONE NORMALE O GAUSSIANA

Definition

The **normal distribution** is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty < x < \infty \qquad \text{(2-21)}$$

The mean of the normal distribution is μ $(-\infty < \mu < \infty)$ and the variance is $\sigma^2 > 0$.

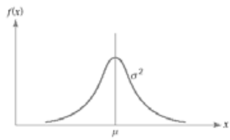


Figure 2-16 The normal distribution.

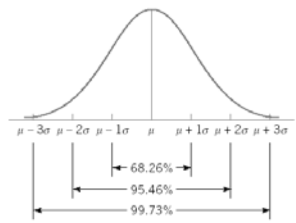


Figure 2-17 Areas under the normal distribution.

DISTRIBUZIONE NORMALE O GAUSSIANA

EXAMPLE 2-7

The tensile strength of paper used to make grocery bags is an important quality characteristic. It is known that the strength—say, x—is normally distributed with mean $\mu = 40$ lb/in^2 and standard deviation $\sigma = 2$ lb/in^2, denoted $x \sim N(40, 2^2)$. The purchaser of the bags requires them to have a strength of at least 35 lb/in^2. The probability that a bag produced from this paper will meet or exceed this specification is $P\{x \geq 35\}$. Note that

$$P\{x \geq 35\} = 1 - P\{x \leq 35\}$$

To evaluate this probability from the standard normal tables, we standardize the point 35 and find

$$P\{x \leq 35\} = P\left\{z \leq \frac{35-40}{2}\right\} = P\{z \leq -2.5\} = \Phi(-2.5) = 0.0062$$

Consequently, the desired probability is

$$P\{x \geq 35\} = 1 - P\{x \leq 35\} = 1 - 0.0062 = 0.9938$$

Figure 2-18 shows the tabulated probability for both the $N(40, 2^2)$ distribution and the standard normal distribution. Note that the shaded area to the left of 35 lb/in^2 in Fig. 2-18 represents the fraction nonconforming or "fallout" produced by the bag manufacturing process.

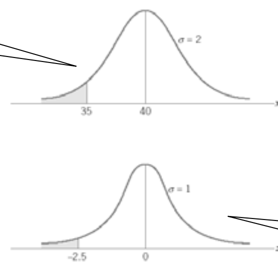
DISTRIBUZIONE NORMALE O GAUSSIANA

Se X è una variabile aleatoria normale con media $E(X) = \mu$ e varianza $V(X) = \sigma^2$, la variabile aleatoria

$$Z = \frac{(X - \mu)}{\sigma}$$

è una variabile aleatoria normale con $E(Z) = 0$ e varianza $V(Z) = 1$, dunque è una variabile aleatoria **normale standard**.

distribuzione normale originale



distribuzione normale standard

Figure 2-18 Calculation of $P\{x \le 35\}$ in Example 2-7.

DISTRIBUZIONE COLLEGATE ALLA NORMALE: χ^2

Sia $X_1, X_2, \ldots, X_n$ un campione casuale estratto da una distribuzione normale con media μ e varianza σ^2 incognite. La quantità

$$X^2 = \frac{(n-1)S^2}{\sigma^2} \tag{4.53}$$

ha una distribuzione chi-quadro con $n - 1$ gradi di libertà, che indichiamo in maniera abbreviata con χ^2_{n-1}. In generale, la funzione di densità di probabilità di una variabile aleatoria chi-quadro è

$$f(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{(k/2)-1} e^{-x/2} \qquad x > 0 \tag{4.54}$$

dove k è il numero di gradi di libertà e $\Gamma(k/2)$ è stata definita nel Paragrafo 4.5.1.

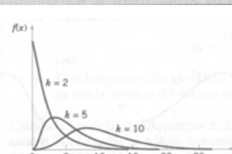


Figura 4.19 Funzioni di densità di probabilità di diverse distribuzioni χ^2.

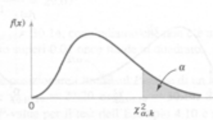


Figura 4.20 Il punto percentuale $\chi^2_{\alpha,k}$ della distribuzione χ^2.

DISTRIBUZIONE COLLEGATE ALLA NORMALE: T di Student

Sia $X_1, X_2, \ldots, X_n$ un campione casuale estratto da una distribuzione normale con media μ e varianza σ^2 incognite. La quantità

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

ha una distribuzione t con $n - 1$ gradi di libertà.

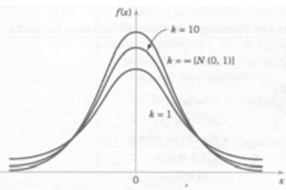


Figura 4.14 Funzioni di densità di probabilità di alcune distribuzioni t.

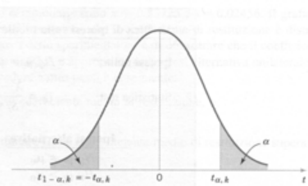


Figura 4.15 Punti percentuali della distribuzione t.

DISTRIBUZIONE COLLEGATE ALLA NORMALE: F

La distribuzione F

Siano W e Y variabili aleatorie chi-quadro indipendenti, rispettivamente con u e v gradi di libertà. Allora il rapporto

$$F = \frac{W/u}{Y/v} \tag{5.18}$$

ha la funzione densità di probabilità

$$f(x) = \frac{\Gamma\left(\frac{u+v}{2}\right)\left(\frac{u}{v}\right)^{u/2} x^{(u/2)-1}}{\Gamma\left(\frac{u}{2}\right)\Gamma\left(\frac{v}{2}\right)\left[\left(\frac{u}{v}\right)x + 1\right]^{(u+v)/2}}, \; 0 < x < \infty \tag{5.19}$$

e si dice che segue la distribuzione F con u gradi di libertà al numeratore e v gradi di libertà al denominatore. Tale distribuzione viene indicata di solito con la notazione abbreviata $F_{u,v}$.

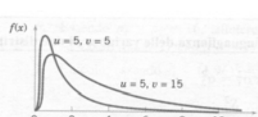


Figura 5.4 Funzione di densità di probabilità di due distribuzioni F.

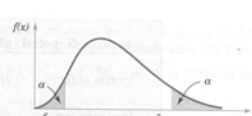


Figura 5.5 Punti percentuali superiore e inferiore della distribuzione F.

DISTRIBUZIONE ESPONENZIALE

Definition

The **exponential distribution** is

$$f(x) = \lambda e^{-\lambda x} \qquad x \geq 0 \tag{2-31}$$

where $\lambda > 0$ is a constant. The **mean** and **variance** of the exponential distribution are

$$\mu = \frac{1}{\lambda} \tag{2-32}$$

and

$$\sigma^2 = \frac{1}{\lambda^2} \tag{2-33}$$

respectively.

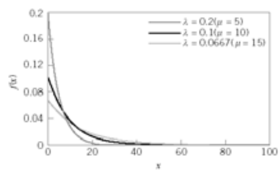


Figure 2-21 Exponential distributions for selected values of λ.

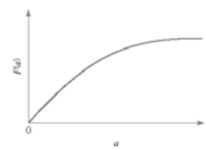


Figure 2-22 The cumulative exponential distribution function.

DISTRIBUZIONE WEIBULL

Definition

The **Weibull distribution** is

$$f(x) = \frac{\beta}{\theta}\left(\frac{x}{\theta}\right)^{\beta-1} \exp\left[-\left(\frac{x}{\theta}\right)^{\beta}\right] \qquad x \geq 0 \tag{2-41}$$

where $\theta > 0$ is the **scale parameter,** and $\beta > 0$ is the **shape parameter.** The **mean** and **variance** of the Weibull distribution are

$$\mu = \theta\Gamma\left(1 + \frac{1}{\beta}\right) \tag{2-42}$$

and

$$\sigma^2 = \theta^2\left[\Gamma\left(1 + \frac{2}{\beta}\right) - \left\{\Gamma\left(1 + \frac{1}{\beta}\right)\right\}^2\right] \tag{2-43}$$

respectively.

DISTRIBUZIONE WEIBULL

Quando β = 1, la distribuzione Weibull si reduce alla distribuzione esponenziale.

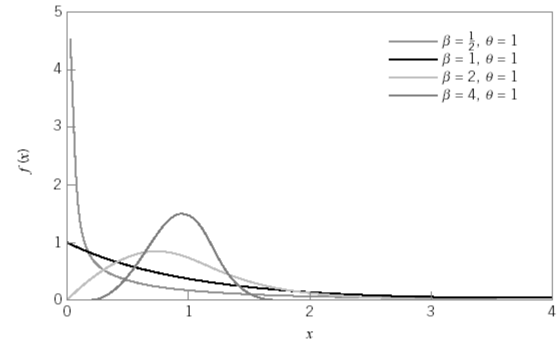


Figure 2-25 Weibull distributions for selected values of the shape parameter β and scale parameter $\theta = 1$.

APPROSSIMAZIONE NORMALE DELLE DISTRIBUZIONI BINOMIALE E POISSON

Se X è una variabile aleatoria binomiale,

$$Z = \frac{X - np}{\sqrt{np(1-p)}} \tag{3.21}$$

è approssimativamente una variabile aleatoria normale standard. Di conseguenza, si possono usare le probabilità calcolate in base a Z per approssimare le probabilità di X.

Se X è una variabile aleatoria di Poisson con $E(X) = \lambda$ e $V(X) = \lambda$,

$$Z = \frac{X - \lambda}{\sqrt{\lambda}} \tag{3.22}$$

è approssimativamente una variabile aleatoria normale standard.