Central Limit Theorem: Implementazione ed Applicazione

1. *Scenario*

Lo scenario è quello della gestione dei resi, processo spesso definito come ‘reverse logistic’, è la gestione degli articoli restituiti dai punti vendita in un centro di distribuzione.
Dopo il ricevimento, i prodotti vengono selezionati, organizzati e controllati per verificarne la qualità. Se sono in buone condizioni, questi prodotti possono essere riforniti in magazzino e aggiunti al conteggio dell'inventario in attesa di essere riordinati.
Quello che si vuole fare in questo lavoro è di usare il teorema del limite centrale per stimare il carico di lavoro (workload) per il processo di gestione dei resi, usando una distribuzione normale basata sulla media e sulla deviazione standard della storia dei records ricevuti.

Si supponga, quindi, di essere l’Inbound Manager di una compagnia multinazionale di vestiti che vende al dettaglio (fast-fashion). Un grande problema che si va a presentare è la mancanza di visibilità del workload per il processo di reso. Infatti, a causa delle limitazioni del sistema, non si riesce ad ottenere un avviso di spedizione anticipato (ASN) prima di aver ricevuto i resi dai negozi.

1) Si ricevono i cartoni tramite pallet che si scaricano dal camion;

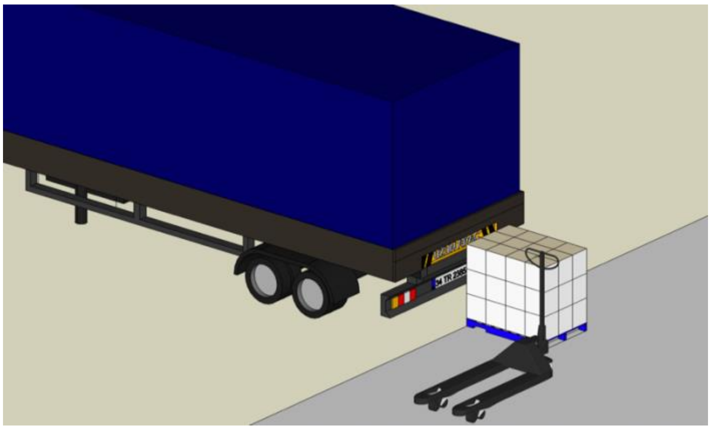


Figura 1: Unloading area (Zona di scarico)

2) Si apre il cartone e si ispezionano gli articoli restituiti.

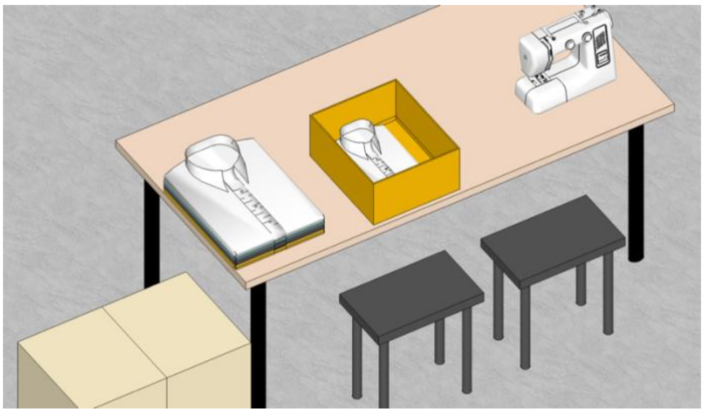


Figura 2: Quality Inspection Workstation (Postazione d’ispezione della qualità)

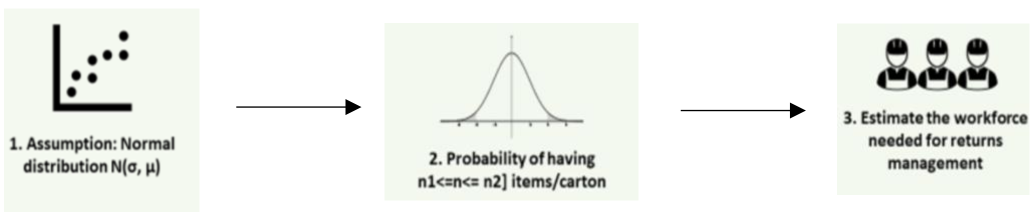
Per ogni capo (camicia, abito…) gli operatori devono eseguire:
- Controllo qualità per garantire che il prodotto possa essere restocked (ridistribuito);
- Relabelling (Rietichettatura);
- Re-packing (Riconfezionamento).

A questo punto, conoscendo la produttività per ogni articolo, si vuole stimare il workload in ore in base al numero di casse che si ricevono in una settimana.

Si suppone che, basandosi sulla storia dei dati negli ultimi 24 mesi, si ha:
- Una media di 23 items per cartone;
- Una deviazione standard di 7 items.

Il team di lavoro è in grado di gestire un massimo di 30 articoli e, quindi, oltre questa soglia è necessario assumere dei lavoratori temporanei per raggiungere l'obiettivo di capacità giornaliera.

La domanda che si pone è la seguente:
Si può stimare la probabilità di avere meno di 30 articoli per cartone che si riceveranno ogni settimana?
Si, usando il Teorema del Limite Centrale.
Il procedimento è caratterizzato dai seguenti steps:



2. *Teorema del Limite Centrale*

Il teorema del limite centrale, definito da qualcuno come il teorema più bello della matematica, stabilisce che, quando si aggiungono variabili casuali indipendenti, la loro somma normalizzata tende a una distribuzione normale anche quando le stesse variabili originali non sono normalmente distribuite.

2.1. Definizione

Si introducono delle notazioni:
- μ, σ : Media e Deviazione standard della popolazione totale;
- $\bar{x}, s$: Media e Deviazione standard dei samples;
- $\bar{X}, S$: Variabili random della media e deviazione standard dei samples della popolazione totale;

Nel contesto esposto, la popolazione totale è l'intera portata dei cartoni ricevuti dai negozi con una media μ = 23 items per cartone e una deviazione standard di σ = 7 items per cartone.
Se si prendono n campioni di cartoni Xn (ad esempio un campione può essere un lotto di cartoni ricevuti in una certa data) si ha:

- $X_1, X_2, \dots, X_n$: sono n random samples tratti da una popolazione con media complessiva (overall mean) μ e una varianza finita σ^2;
- $\bar{X}_n$: Media del sample n;
- $Z = \lim_{n\to\infty} \sqrt{n} \left(\frac{\bar{X}_n - \mu}{\sigma} \right)$: È la distribuzione standard normale.

In altre parole, se si misura in modo casuale il numero di articoli per cartone utilizzando n samples e si assume che le osservazioni siano indipendenti e identicamente distribuite (i.i.d.), la distribuzione di probabilità del sample medio si avvicinerà strettamente a una distribuzione normale.

Nota: per garantire che le osservazioni siano indipendenti e identicamente distribuite, si assume che i samples siano costruiti sulla base di batch di resi provenienti da tutti i negozi in un ambito che copre il 100% dello SKU (Stock Keeping Unit) attivo.

2.2. Applicazione

Quindi, si va ad assumere che il numero medio di items per scatola segua una distribuzione normale con media $\mu = 23$ items per cartone e una deviazione standard di $\sigma = 7$ items per cartone.

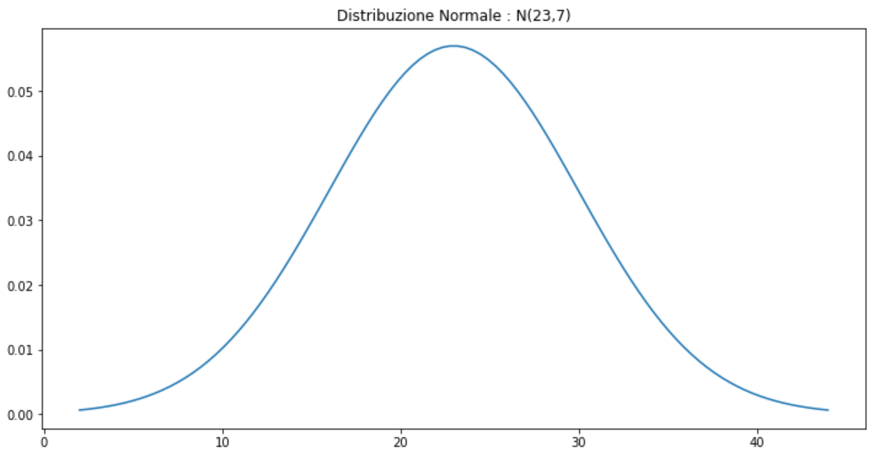


Figura 3: Distribuzione Normale della popolazione

A questo punto ci si chiede qual'è la probabilità di avere meno di 30 items per cartone.

2.3. Probabilità ‹30 items per cartone

```
Probabilità di avere 30 items/cartone è 84.13%
```

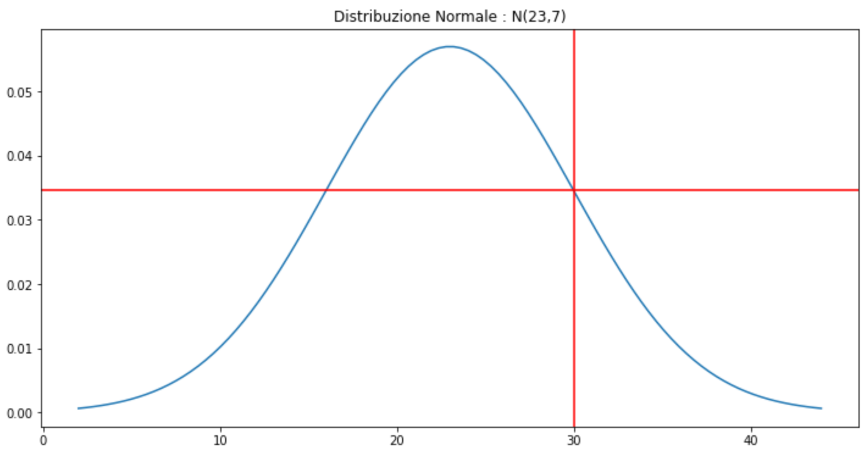


Figura 4: Distribuzione Normale della popolazione

Codice Python:

```python
# Parametri
mu = 23
sigma = 7

# Set up della funzione di probabilità
stats.norm(mu, sigma)

# Probabilità X = 30
x1 = 30

# Funzione di Densità Cumulativa
y1 = stats.norm(mu, sigma).cdf(x1)

# Probabiltà della Funzione Densità
z1 = stats.norm(mu, sigma).pdf(x1)

plt.figure(figsize=(12, 6))
ax = plt.gca()
x = np.linspace(mu - 3*sigma, mu + 3*sigma, 100)
plt.plot(x, stats.norm.pdf(x, mu, sigma))
plt.axvline(x1, color='red')
plt.axhline(z1, color='red')
plt.title('Distribuzione Normale : N(23,7)')
print("Probabilità di avere {} items/cartone è {}%".format(x1, round(100*y1,2)))
plt.show()
```

2.4. 95% di probabilità di avere meno di k items per scatola

Si suppone che l'obiettivo KPI (Key Performace Indicator) sia che almeno il 95% dei resi venga elaborato lo stesso giorno.

A questo punto ci si chiede quanti items si devono assumere per fare in modo che il team riesca a gestire il 95% del workload previsto.

```
Si ha il 95% di probabilità che X ⩽ 34.514 items/cartone
```

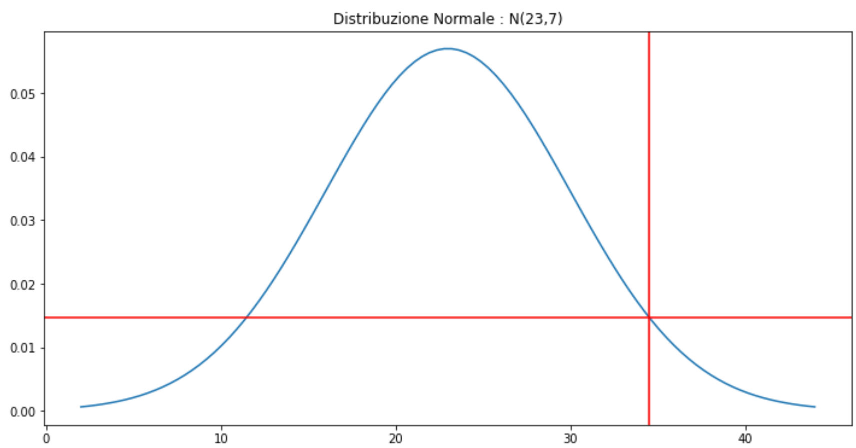


Figura 5: Distribuzione Normale della popolazione

Codice Python:

```python
# Parametri
mu = 23
sigma = 7

# Set up della funzione di probabilità
stats.norm(mu, sigma)

# Probabilità X = 30
y1 = 0.95

# Percent point function (inversa della funzione di densità cumulativa)
x1 = stats.norm(mu, sigma).ppf(y1)

# Probablità della funzione di densità
z1 = stats.norm(mu, sigma).pdf(x1)

plt.figure(figsize=(12, 6))
ax = plt.gca()
x = np.linspace(mu - 3*sigma, mu + 3*sigma, 100)
plt.plot(x, stats.norm.pdf(x, mu, sigma))
plt.axvline(x1, color='red')
plt.axhline(z1, color='red')
plt.title('Distribuzione Normale : N(23,7)')
print("Si ha il 95% di probabilità che X <= {} items/cartone".format(round(x1,3)))
plt.show()
```

Si evince che, se si ridimensiona il team di lavoro in base a 35 articoli per cartone, si raggiungerà in media il 95% dell'obiettivo.

Si riesce, attraverso questa metodologia, a ridimensionare il team di lavoro sulla base di ipotesi supportate da potenti strumenti statistici. In questo modo, ad esempio in un business in evoluzione (maggior numero di collezioni, aperture di nuovi negozi, etc.), eseguendo più volte l'anno questa analisi si va a far fronte alla variabilità in input che, chiaramente, contraddistinguerà il sistema.