

Relazione Simulazione

Simulatore casse ipercoop Collestrada

Matteo Capparrucci

May 12, 2016

Contents

1	Osservazione del sistema reale	2
1.1	Acquisizione dei Dati	2
1.2	Convalida Osservazioni Sperimentali	3
2	Studio Analitico	5
3	Simulatore	6
3.1	Prove Ripetute	7
3.2	Stabilizzazione tramite media delle medie	8
3.3	Stabilizzazione tramite media teorica	9
4	Conclusioni	10

Testo

Il gestore del vostro supermercato (IPERCOOP Collestrada) ha problemi per determinare il numero di cassiere da avere il venerdì pomeriggio, uno dei periodi di maggiore affluenza. Poiché la tua compagna si lagna spesso delle attese alla cassa e ti ha sfidato a mettere in pratica la tua preparazione sulla teoria delle code, tu ti incarichi del compito di aiutare il gestore. Decidi quindi – a causa della complessità di decisione e di diniego, dell’abilità di aggiungere o togliere cassiere o casse “fai da te” (che incrementano o decrementano il μ di canale) e altre complessità – di simulare il modello di sistema. Sviluppa il

modello di simulazione e implementalo nel linguaggio che tu desideri. Tenta di stimare, a partire da tue reali osservazioni su alcuni sabati mattina, i modelli di arrivo e di servizio, così come la disciplina delle code. Usa parte delle osservazioni per sviluppare le distribuzioni empiriche e convalida poi il simulatore usando le rimanenti osservazioni (al 90% del livello di confidenza). Quindi cerca di determinare le soluzioni per i problemi del gestore, tenendo presente anche delle casse “fai da te”.

1 Osservazione del sistema reale

Il centro commerciale di Collestrada è uno dei maggiori supermercati della provincia di Perugia e, come tale, è stato progettato per gestire un grande flusso di clienti. Nello specifico la configurazione delle casse offre la possibilità di far funzionare 18 casse Normali, 6 Casse “fai da te” e 6 cassi “Salva Tempo”. Nel progetto osserveremo esclusivamente i due gruppi di casse: “normali” e “fai da te” comparando gli effetti dell’aumento o della diminuzione della congestione del traffico alle casse al variare dei numeri di queste.

1.1 Acquisizione dei Dati

Le osservazioni sono State eseguite presso il centro commerciale Collestrada in data 22/01/2016 in un intervallo di tempo compreso tra le ore 19:30 e le ore 22:00 tramite l’uso di una semplice applicazione android progettata “ad hoc” per consentire all’osservatore di ottenere dati precisi anche nel caso di arrivi ravvicinati e numerosi presso le code. Osservando rispettivamente l’unica coda comune per le casse “fai da te” e una coda campione, tra le 6 casse normali aperte durante la raccolta dei dati, abbiamo ottenuto le seguenti osservazioni:



Figure 1: App Android

Cassa	# TA	# TS	Mean(TA)	Mean(TS)
Normale	73	71	62.8	86.0
Fai da te	67	14	51.9	91.8

1.2 Convalida Osservazioni Sperimentali

Al fine di realizzare un Simulatore che descrivesse il comportamento reale del sistema si è quindi proceduto con la convalida dei dati al fine di individuare le distribuzioni di probabilità che meglio potevano rappresentare i tempi di arrivo degli utenti nel sistema e i vari tempi di servizio.

Table 1: Goodness of Fit Tempi di Arrivo Cassa Normale (Esponenziale)

Intervalli		f_i	$p(i)$	F_i	$\frac{(f_i - F_i)^2}{F_i}$
0	11	12	0,147532756	8,999498117	1,000390404
11	20	10	0,1043675881	6,366422873	2,073830627
20	32	11	0,1195564621	7,292944189	1,884323043
32	45	11	0,1080575278	6,591509193	2,948458482
45	60	7	0,1018093781	6,210372065	0,100398538
60	105	10	0,2007612813	12,24643816	0,4120777282
105	$+\infty$	12	0,2179150066	13,2928154	0,1257349636
		61	1	61	8,545213786

$$\bar{x} = 62.8 \quad \sigma = 75.0$$

Table 2: Goodness of Fit Tempi di Servizio Cassa Normale (Esponenziale)

Intervalli		f_i	$p(i)$	F_i	$\frac{(f_i - F_i)^2}{F_i}$
0	40	22	0,3763121989	22,95504414	0,0397345915
40	80	18	0,2347013279	14,316781	0,9475665096
80	$+\infty$	21	0,3889864732	23,72817486	0,3136751196
		61	1	61	1,300976221

$$\bar{x} = 86.08 \quad \sigma = 83.4$$

Table 3: Goodness of Fit Tempi di Arrivo Cassa “Fai Da Te” (Esponenziale)

Intervalli		f_i	$p(i)$	F_i	$\frac{(f_i - F_i)^2}{F_i}$
0	10	10	0,1662764104	11,14051949	0,1167615852
10	19	10	0,1258723931	8,43345034	0,2909933346
19	30	10	0,128334078	8,598383227	0,2284766248
30	40	9	0,0963600262	6,456121756	1,002353544
40	60	10	0,1473170017	9,870239111	0,001705925054
60	100	9	0,173576569	11,62963012	0,5945979794
100	inf	9	0,1622635217	10,87165595	0,3222228531
		67	1	67	2,557111846

$$\bar{x} = 51.9 \quad \sigma = 58.0$$

Table 4: Kolmogorov-Smirnov Tempi Arrivo di Cassa “Fai Da Te” (Esponenziale)

Intervalli		f_i	$p(i)$	$f(i)$	d_i	D_i	$ d_i - D_i $
0	40	5	0,59854288	0,3575	0,35714	0,59854288	0,2414
40	60	4	0,1470913036	0,2857	0,64286	0,7456341836	0,1028
60	inf	5	0,2543658164	0,3571	1	1	0
		14	1	1			

$$\bar{x} = 56.8 \quad \sigma = 43.8$$

Come visibile nelle tabelle (1 e 2), che descrivono i risultati del metodo Goodness of fit, i tempi di interarrivo e di servizio delle casse normali sono rappresentati da delle distribuzioni esponenziali che ricadono nell'intervallo tra il P_{10} percentile e il P_{90} della distribuzione X^2 . Lo stesso avviene per i tempi di interarrivo per le casse “Fai Da Te” (3 e 4), mentre i tempi di servizio di queste ultime risultano essere anch'essi esponenziali ma sono stati convalidati mediante il test Kolmogorov Smirnov a causa dello scarso numero di dati osservati. In ogni caso, il massimo scarto riscontrato tra la cumulativa osservata e la teorica risulta minore del valore critico di Kolmogorov-Smirnov $D_{0,10}$ per il rispettivo valore n e quindi il risultato è accettabile.

2 Studio Analitico

Convalidate le distribuzioni possiamo definire il modello teorico del sistema di casse del centro commerciale. Nello specifico otterremo una configurazione con n M/M/1 in parallelo (per le casse normali) e una M/M/6 (per le casse “Fai Da Te”) senza la possibilità di cambio di coda.

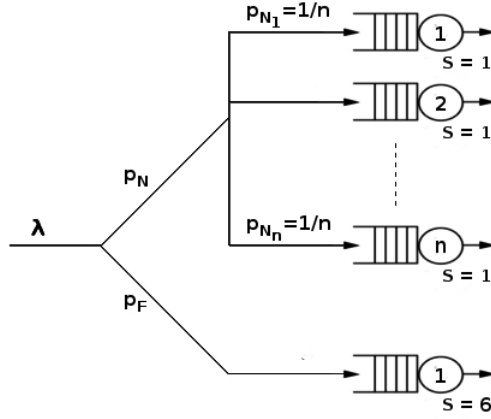


Figure 2: Schema del sistema di code

Fissato il sistema nella configurazione osservata durante la raccolta dei dati, definiamo per prima cosa $n = 6$ (quindi limitiamo il numero delle casse normali aperte a 6 come durante il periodo osservato) e di conseguenza ricaviamo il numero di arrivi al secondo per l'intero sistema e quindi quello verso le due tipologie di casse:

$$\lambda = \lambda_F + \lambda_N = \frac{1}{Ta_F} + \frac{n}{Ta_{N_i}} = \frac{68.9 + 6 \cdot 55}{55 \cdot 68.9} = 0.105s^{-1}$$

dove:

$$\lambda_F = \frac{1}{Ta_F}$$

$$\lambda_N = \sum_{i=1}^s \lambda_{N_i} \cdot p_i = \lambda_{N_i} \cdot n \text{ dato che } p_n = \frac{1}{n} \forall i \in \{1..n\}$$

$$\lambda_{N_i} = \frac{1}{Ta_N} \text{ tc : } i \in \{1..n\}$$

$$Ta_N = 68.9s$$

$$Ta_F = 55.0s$$

Da queste è facile ricavare le probabilità p_N , p_F come:

$$\begin{aligned} p_N &= \lambda_N / \lambda = 83\% \\ p_F &= \lambda_F / \lambda = 17\% \end{aligned}$$

A questo punto è facile calcolare l'indice di congestione ρ delle due tipologie di coda e notiamo che se da un lato la coda di tipo M/M/6 ha ρ_F molto basso, dunque la probabilità di fare coda utilizzando le casse “Fai Da Te” è pressoché nulla, il sistema non è invece da considerarsi stazionario nel caso delle casse normali. Queste ultime infatti presentano un $\rho_N = 1.23$ con $n = 6$ e $\rho_N = 1.05$ con $n = 7$. Definiamo il sistema stazionario quindi solo per $n \geq 8$ ad esempio nel caso di $\rho_N(8) = 0.92$. Si decide quindi di porre l'attenzione del progetto sulle attese medie in coda degli utenti delle due tipologie di casse che descriviamo come:

$$\begin{aligned} T_{W_N} &= \frac{\rho_N / \mu}{1 - \rho} \\ T_{W_F} &= \frac{\pi_n}{n\mu(1 - \rho_F)^2} \end{aligned}$$

E definiamo il tempo medio di attesa generale nel sistema come:

$$\begin{aligned} T_W &= T_{W_N} \cdot p_N + T_{W_F} \cdot p_F \\ \text{Con } T_W &= 8.39s \text{ per } n = 8 \text{ e } T_{W_N} = 1011.0397s \text{ e } T_{W_F} = 0.0014 \end{aligned}$$

3 Simulatore

Il simulatore del sistema in Fig. 2 è stato realizzato utilizzando Python 2.7.10 e la libreria Simpy¹ che permette di implementare e gestire semplicemente dei simulatori ad eventi discreti. Nello specifico, il simulatore procederà nella sua esecuzione schedulando un istante di tempo alla volta (settato per essere un secondo nel mondo reale) e permettendo il controllo del numero di “clienti” da far transitare nel sistema per ogni esecuzione.

¹<https://simpy.readthedocs.org/en/latest/>

```
#####
##### Simulazione #####
Tempo di Simulazione = 3601
Numero Casse Normali Aperte = 6
Numero Casse "Fai da te" Aperte = 6
Numero clienti serviti = 268
Numero clienti in coda = 91
Tempo di Attesa Medio = 357.854927286
Tempo di Servizio Medio = 78.3716553987
Tempo di Arrivo Medio = 10.0525006785
#####
Warning rho > 1 Sistema non stabile
```

Figure 3: Singola Esecuzione in condizioni reali

3.1 Prove Ripetute

Al fine di poter utilizzare i dati forniti dal simulatore occorre per prima cosa valutare la stazionarietà del sistema. Questa tuttavia richiede che, oltre ad avere dei valori di media e varianza stazionari, si abbiano osservazioni x_1, \dots, x_n statisticamente indipendenti. Questo, tuttavia, non è sempre possibile inquanto, come nel nostro caso, talvolta si scelgono come parametri di riferimento dei valori per loro natura auto-correlati come ad esempio i tempi di attesa. Questi infatti intuitivamente non dipendono solo dalla frequenza di arrivo degli utenti nel sistema ma anche dallo stato stesso di congestione della coda, e quindi dagli altri tempi di attesa. Tuttavia la distribuzione delle medie $\bar{x}(n) = n^{-1} \sum_{i=1}^n x_i$ per n sufficientemente grande, con $E(\bar{x}) = \eta$, si comporta come quella di dati indipendenti. Con il metodo delle prove ripetute quindi verranno eseguiti p differenti esperimenti, ognuno dei quali composto da n_j osservazioni ($n_j \in [50 \dots 100]$). Ognuno dei p esperimenti sarà eseguito dopo aver stabilizzato il sistema, ovvero dopo aver eliminato l'influenza dello stato iniziale. Dato che assumiamo che alla fine di un esperimento p lo stato del sistema sia ancora stabile potremo eseguire ulteriori esperimenti subito dopo la conclusione di p .

Per stimare la condizione di stabilità del sistema, utilizzeremo il valore teorico η ovvero la media teorica ottenuta dallo studio analitico del sistema, e quindi $E(\bar{x}) = \eta$ o in alternativa gli stimatori di Gordon: $e(n) = p^{-1} \sum_{i=1}^p \bar{x}_j(n)$ e $s^2(\bar{x}) = (p-1)^{-1} \sum_{j=1}^p [\bar{x}_j(n) - e(n)]^2$ con media campionaria pari $\bar{x}_j(n) = n^{-1} \sum_{i=1}^n x_{ij}$.

3.2 Stabilizzazione tramite media delle medie

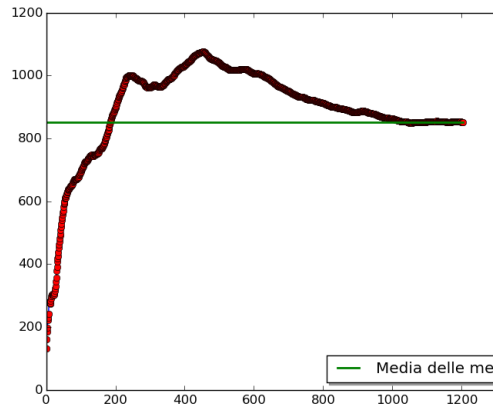


Figure 4: Stabilizzazione

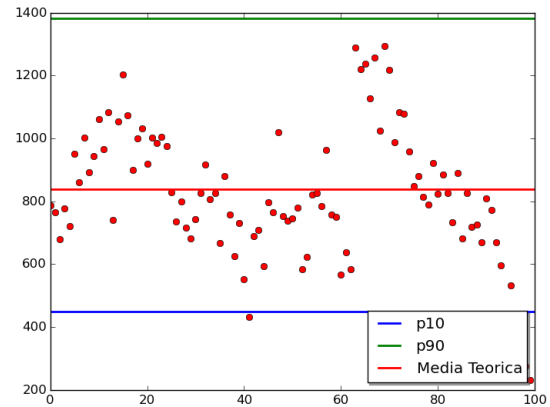


Figure 5: Distribuzione esperimenti

```
#####
##### Dati Teorici #####
lambda = 0.105
lambda_mml = 0.01089375
lambda_mmm = 0.01785
mu_MML = 0.0118063754427
mu_MMm = 0.0228310502283
rho_MML = 0.922700625
rho_MMm = 0.130305
T_W_MML = 1011.03977797
T_W_MMm = 0.00140238840996
T_W_tot = 839.163254124
n = 8
m = 6
#####
##### Prove Ripetute #####
Tempo totale Simulazione = 883270
Tempo di Simulazione = 0
Numero Casse Normali Aperte = 8
Numero Casse "Fai da te" Aperte = 6
Numero clienti = 92700
Numero clienti in coda = 40
Esecuzioni totali = 1205
Numero esecuzioni p per la stazionarieta' del sistema = 1105
Valore Stazionario = 851.015261132
Tempo di Attesa medio = 892.317890338
p10 = 448.880536925
p90 = 1381.66219735
Percentuale = 94
#####
```

Figure 6: Output Simulatore

3.3 Stabilizzazione tramite media teorica

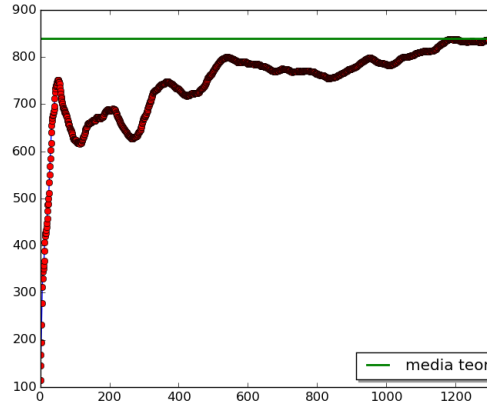


Figure 7: Stabilizzazione

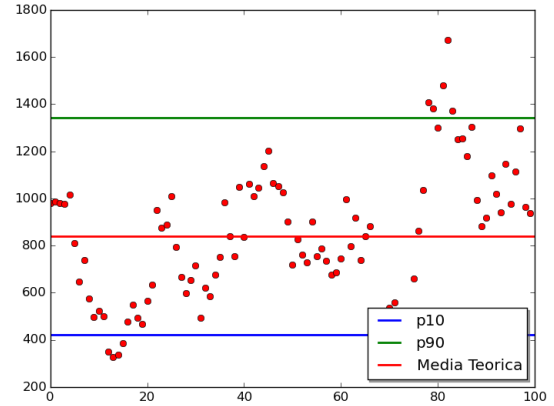


Figure 8: Distribuzione esperimenti

```
#####
##### Dati Teorici #####
lambda = 0.105
lambda_mml = 0.01089375
lambda_mmm = 0.01785
mu_MML = 0.0118063754427
mu_MMm = 0.0228310502283
rho_MML = 0.922700625
rho_MMm = 0.130305
T_W_MML = 1011.03977797
T_W_MMm = 0.00140238840996
T_W_tot = 839.163254124
n = 8
m = 6
#####
##### Prove Ripetute #####
Tempo totale Simulazione = 944852
Tempo di Simulazione = 22712
Numero Casse Normali Aperte = 8
Numero Casse "Fai da te" Aperte = 6
Numero clienti = 99316
Numero clienti in coda = 99
Esecuzioni totali = 1298
Numero esecuzioni p per la stazionarieta' del sistema = 1198
Valore Stazionario = 836.747147657
Tempo di Attesa medio = 746.603349894
p10 = 422.161833788
p90 = 1341.45983474
Percentuale = 90
#####
```

Figure 9: Output Simulatore

4 Conclusioni

In conclusione possiamo affermare che il simulatore implementato rappresenta in modo accettabile il sistema reale da noi osservato (Fig. 3). Inoltre troviamo che il raggiungimento dello stato di stabilizzazione può essere effettivamente raggiunto con successo sia attraverso l'utilizzo dello stimatore della media campionaria sia utilizzando la media teorica. Nel secondo caso però, è richiesto lo scarto di un tempo di simulazione più lungo per il raggiungimento della stabilità.

Per quanto riguarda il sistema reale (ipercoop collestrada) appare evidente come le casse “fai da te” introducano nel sistema una variazione del tempo di attesa praticamente trascurabile a causa del loro bassissimo tasso di utilizzo ($\rho_F = 0.13$).

Per quanto riguarda le casse normali invece notiamo come il sistema di partenza non sia stazionario se non con l'apertura di almeno altre due casse, sebbene con un tempo di attesa medio di più di 15 minuti. Al fine di abbattere affinché diventi minore di 5 minuti (un tempo di attesa che consideriamo accettabile) occorre aumentare il numero di casse fino almeno ad 11 ottenendo un tempo di attesa media nel sistema pari a ≈ 2.5 min.

```
#####
##### Dati Teorici #####
lambda = 0.105
lambda_mml = 0.00792272727273
lambda_mmm = 0.01785
mu_MML = 0.0118063754427
mu_MMm = 0.0228310502283
rho_MML = 0.671055
rho_MMm = 0.130305
T_W_MML = 172.789853927
T_W_MMm = 0.00140238840996
T_W_tot = 143.415817165
n = 11
m = 6
#####
##### Prove Ripetute #####
Tempo totale Simulazione = 446812
Tempo di Simulazione = 0
Numero Casse Normali Aperte = 11
Numero Casse "Fai da te" Aperte = 6
Numero clienti = 47011
Numero clienti in coda = 19
Esecuzioni totali = 590
Numero esecuzioni p per la stazionarieta' del sistema = 490
Valore Stazionario = 148.783711826
Tempo di Attesa medio = 150.776430559
p10 = 73.0719727954
p90 = 256.508239637
Percentuale = 96
#####
```

Figure 10: Output Simulatore con 11 casse normali

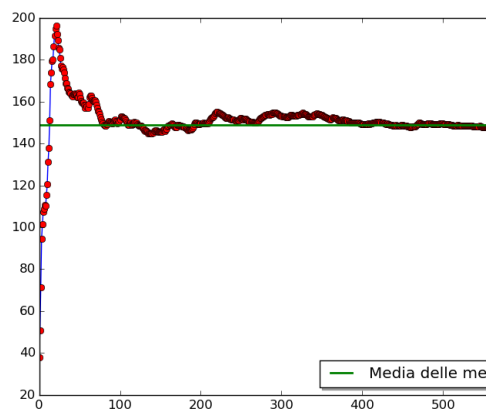


Figure 11: Stabilizzazione

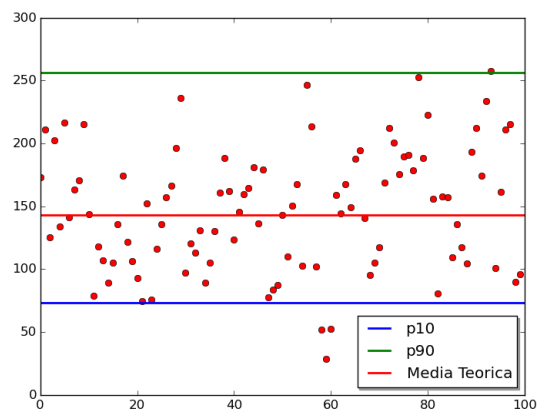


Figure 12: Distribuzione esperimenti