distribuzioni di probabilità diverse modellano caratteristiche diverse.
Sono degli esempi, ma con la teoria dietro, dovremmo acquisire sensibilità al problema.

Dicii

Performance Modeling
of Computer Systems and Networks

Prof. Vittoria de Nitto Personè

Discrete Random Variates: applications

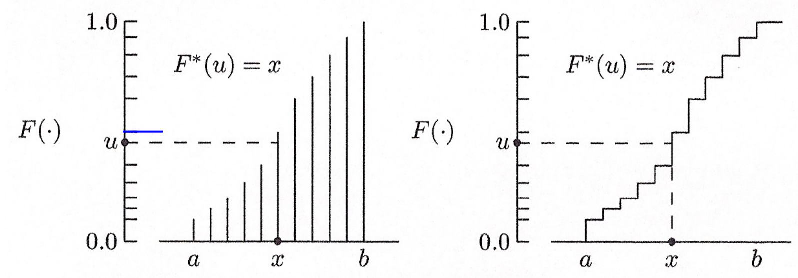
Università degli studi di Roma Tor Vergata
Department of Civil Engineering and Computer Science Engineering

1

cumulativa discreta equilikely (corrispettivo discreto dell'uniforme).
probabilità tutte uguali, sia continue che discrete.

Equilikely(a,b) $F(x) = \dfrac{x-a+1}{b-a+1} \qquad x = a, a+1, \ldots, b$

$$F^*(u) = \min_x \{x : u < F(x)\}$$



$F(x\text{-}1) \le u < F(x)$ $\longleftrightarrow$ $\dfrac{(x-1)-a+1}{b-a+1} \le u < \dfrac{x-a+1}{b-a+1}$

by Lehmer

da trovare

2

in caso di funzioni semplici la trasformazione è semplice
linea blu = F(x) vero, tutti i valori da u a linea blu esclusa vengono rappresentati da stesso "valore", cioè la 'x' sull'asse delle x.

Discrete Simulation
Discrete RV applications

$$\frac{(x-1)-a+1}{b-a+1} \le u < \frac{x-a+1}{b-a+1}$$

$$x-a \le (b-a+1)u < x-a+1$$

$$x \le a+(b-a+1)u < x+1$$

$$x = a+\lfloor (b-a+1)u \rfloor$$

$$F^*(u) = a+\lfloor (b-a+1)u \rfloor = x$$

Prof. Vittoria de Nitto Personè 3

3

usavamo equilikely per generare un numero(10,50) = numero di domande per articolo nella settimana. aver scelto equilikely significa aver scelto modello slide dopo.

model 1

Discrete Simulation
Discrete RV applications

Example: Inventory System

in program sis2, the demand per time interval is an *Equilikely*(10,50) random variate

```
long Equilikely(long a, long b)
{  return (a + (long) ((b - a + 1) *
Random()));}

long GetDemand(void)
{
  return (Equilikely(10, 50)); }
…
 while (index < STOP) {
 index++;
 …
 inventory -= demand;}
```

Prof. Vittoria de Nitto Personè 4

4

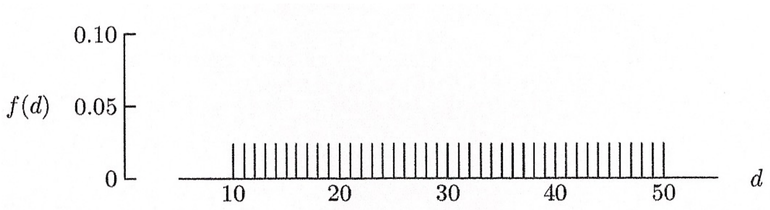
tutti stessa probabilità. equilikely media 30 e deviazione std circa 11.8 (dispersione piccola perchè la funzione è piatta, tutte stessa probabilità.). Non è realistica però, perchè casi limiti dovrebbero essere meno probabili.

model 1

Discrete Simulation
Discrete RV applications

Example: Inventory System

- $\mu=30$, $\sigma = \sqrt{140} \cong 11.8$, and the demand pdf is flat



- this model is not very realistic

Prof. Vittoria de Nitto Personè 5

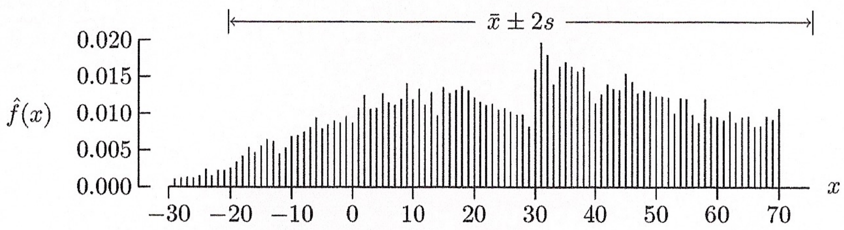
5

questa variazione di sis2 permette di ottenere gli istogrammi. con molti dati posso avere di più di medie e varianze, posso avere proprio la distribuzione!

DE simulation
Discrete–Data Histograms

Example 3.1

- Using sis2mod to generate 10000 weeks of sample data, the inventory level histogram from sis2.out can be constructed (s,S)=(20,80)



- *x* denotes the inventory level prior to review
- $\bar{x} = 27.63, \quad s = 23.98$
- about 98.5% of the data falls within $\bar{x} \pm 2s$

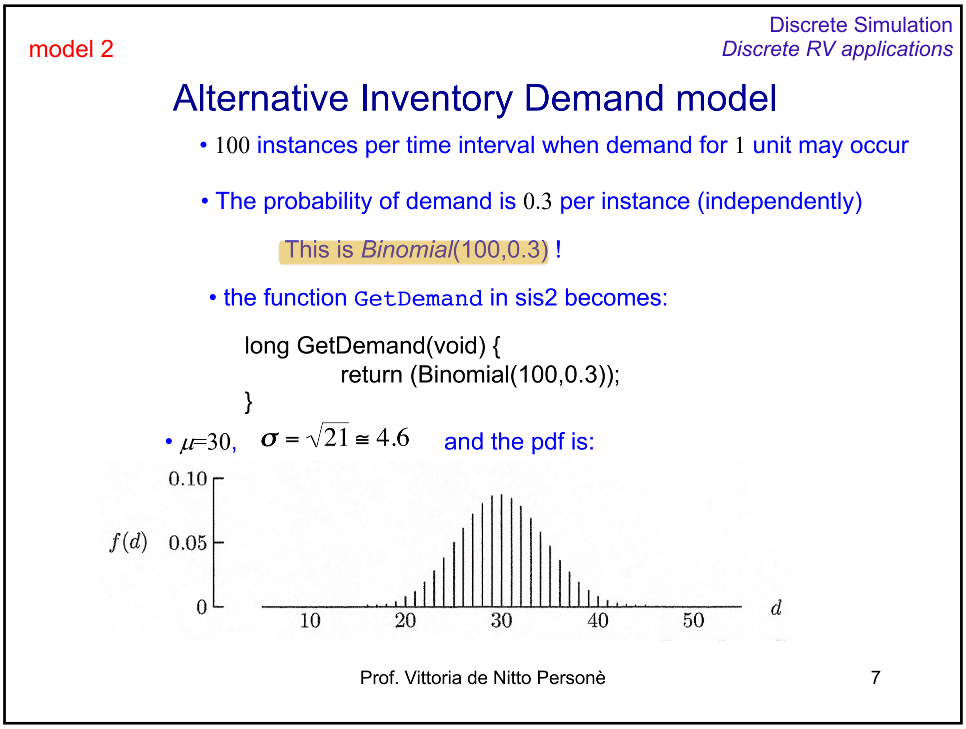
Equilikely(10, 50) ?

Prof. Vittoria de Nitto Personè 6

6

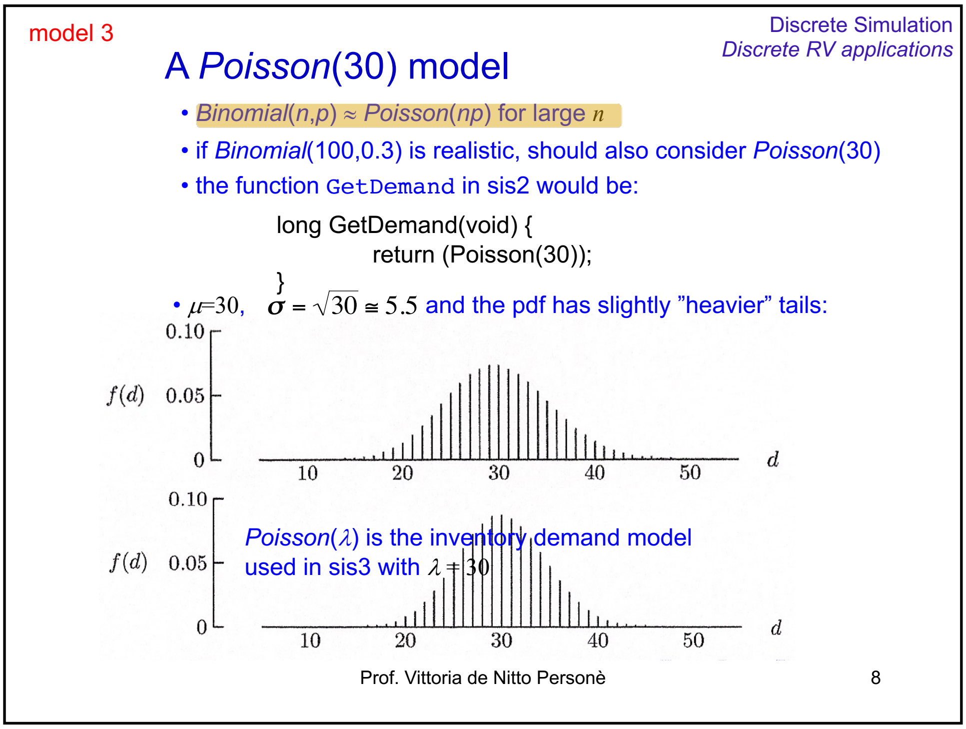
vale disuguaglianza di Chebyschev, ho infatti preso (x-+2s)
Quanto è realistica la equilikely? vedo altri modelli.

provo model2 più articolato: invece di prendere numero da spalmare uniformemente nella settimana, suppongo di prendere per ogni settimana 100 istanze POSSIBILI di domande per articolo, con probabilità di istanza = 0.3 (cosi media 30 come equilikely).
costruendo l'istogramma ottengo la figura: deviazione std ridotta: code scendono (casi limite meno probabili), sicuro è meglio della equilikely.



7

se n grande, posso passare a Poisson di media 30. l'istogramma presenta più varianza e code leggermente più grande. questo modello è usato in sis3.



8

passo a 50 istanze, lasciando l'ipotesi di quantità unitaria per ogni istanza.
(non piu articolo unitario)
ora è possibile ordinare più merci insieme. la domanda è intera con geometrica.

06/05/21

model 4

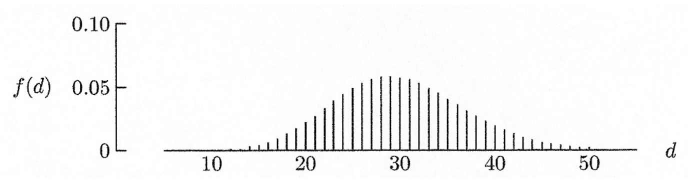
A *Pascal*(50,0.375) model

così media sempre 30

- 50 instances per time interval
- the demand per instance is *Geometric*(*p*) with p=0.375
- the function `GetDemand` in sis2 would be:

```
long GetDemand(void) {
        return (Pascal(50,0.375));
}
```

- μ=30, $\sigma = \sqrt{48} \cong 6.9$ and the pdf has heavier tails than the *Poisson*(30) pdf:



9

se invece di 100 o 50 istanze usassi un valore random?
genero numero istanze da poisson, con troncamento (sennò potrei non avere domande).

model 5

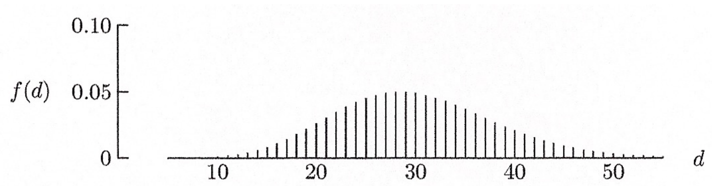
A compound demand model

- the number of demand instances per time interval is *Poisson*(50)
- the demand per instance is *Geometric*(*p*) with p=0.375
- the function `GetDemand` in sis2 becomes:

```
long GetDemand(void) {
        long instances = Poisson(50.0);  //must truncate to avoid 0
        return (Pascal(instances, 0.375));
}
```

random

- μ=30, $\sigma = \sqrt{66} \cong 8.1$ and the pdf has heavier tails



10

varianza aumenta, media rimane stabile.

se volessi la densità? definisco due variabili random: quantità per singola instanza D ed I numero di istanze per unità di tempo

06/05/21

model 5

Discrete Simulation
Discrete RV applications

pdf for the compound demand

- define random variables
  - D: the demand amount
  - I: the number of demand instances per time interval

$$f(d) = \Pr(D = d) = \sum_{i=0}^{\infty} \Pr(I = i)\Pr(D = d \mid I = i)$$

d domande

dato che I = i ...

poiché random non sono più certo!

$d = 0, 1, 2, \ldots$

ho d domande

- to compute $f(d)$, truncate infinite sum: $0 < a \leq i \leq b$

```
/* use the library rvms */
double sum = 0.0;
for (i = a; i <= b; i++)
        sum += pdfPoisson(50.0,i) * pdfPascal(i,0.375,d);
return sum;
/* sum is f(d) */
```

punto "i"

punto "d"

Prof. Vittoria de Nitto Personè 11

11

faccio confronto tra modelli visti (sono tutti modelli di sis2, non guidata dal tempo ma interazione tra processi) e modelli sis3 (next-event) con tempi interdomanda con exp.

Next-Event simulation
InvSys

Comparison of Demand Models

sis2: used an *aggregate* demand for each time interval, generated as an *Equilikely*(10,50) random variate

- Aggregate demand per time interval is random
- Within an interval, time between demand instances is constant
- Example: if aggregate demand is 25, inter-demand time is 1/25=0.04

- Now (sis3) using *Exponential*($1/\lambda$) inter-demand times
  - Demand is modeled as an arrival process
  - Average demand per time interval is λ

Sis3

```
double GetDemand(void)/* ------------------------------------------------- *
                       generate the next demand instance (time) with rate 30 *
                       per time interval and exactly one unit of demand per  *
                       demand instance  * ---------------------------------*/
{ static double time = START;
  SelectStream(0);
  time += Exponential(1.0 / 30.0);
return (time);}
```

Prof. Vittoria de Nitto Personè 12

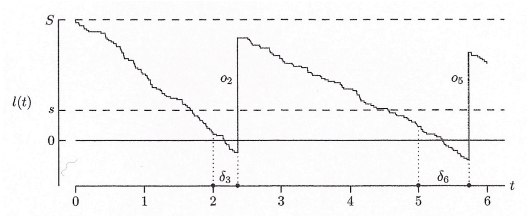
12

altro modello, sis4 basato su sis3 (interdomanda exp). quando l'istanza della domanda c'è o meno è random di probabilità p. per chiedere più di 1 articolo, il numero medio è modellato con geometrica. 06/05/21

model 6

Discrete Simulation
Discrete RV applications

sis3



Program sis4

- Based on sis3 but with a more realistic inventory demand model
- The inter-demand time is an *Exponential*$(1/\lambda)$ random variate
- Whether or not a demand occurs at demand instances is random with probability p
- To allow for the possibility of more than 1 unit of demand, the demand amount is a *Geometric*(p) random
- Expected demand per time interval is $\frac{\lambda p}{(1-p)}$

Prof. Vittoria de Nitto Personè 13

13

ho lambda clienti per settimana, che comprano 0,...,n macchine secondo distr. geometrica.

model 6

Discrete Simulation
Discrete RV applications

The auto dealership

- The inventory demand model for sis4 corresponds to λ customers per week on average
- Each customer will buy
  - 0 autos with prob $(1\text{-}p)$
  - 1 auto with prob $(1\text{-}p)p$
  - 2 autos con prob $(1\text{-}p)p^2$, etc.
- with λ=120 p=0.2, average demand is 30

Somma pesata

$$30.0 = \frac{\lambda p}{(1-p)} = \lambda \sum_{x=0}^{\infty} x(1-p)p^x = \underbrace{\lambda(1-p)p}_{19.2} + \underbrace{2\lambda(1-p)p^2}_{7.68} + \underbrace{3\lambda(1-p)p^3}_{2.304} + \ldots$$

- $\lambda(1\text{-}p)$=96.0 customers buy 0 autos
- $\lambda(1\text{-}p)p$=19.2 customers buy 1 auto
- $\lambda(1\text{-}p)p^2$=3.84 customers buy 2 autos
- $\lambda(1\text{-}p)p^3$=0.768 customers buy 3 autos, etc.

Prof. Vittoria de Nitto Personè 14

14

in media 96 clienti comprano 0 macchine, 19.2 comprano un auto etc...

qui c'è implementazione sis4, con due stream: uno per tempo interdomanda, e uno per le domande. due stream per avere processi disaccoppiati.
getDemand genera quantità intera, e l'inventario viene decretato di quella quantità, mentre in sis3 decrementavo di 1 soltanto (poichè richieste singole).

model 6 sis4.c

uncoupled processes

```c
double GetDemand(long *amount)
{
 static double time = START;
 SelectStream(0);
 time += Exponential(1.0 / 120.0);  /* demand instance */
 SelectStream(2);
 *amount = Geometric(0.2);          /* demand amount   */
 return (time);
}
...
if (t.current == t.demand) { /* process an inventory demand */
   sum.demand += amount;
   inventory  -= amount;
   t.demand    = GetDemand(&amount);
}
```

```c
if (t.current == t.demand) { /* process an inventory demand */
    sum.demand++;
    inventory--;
    t.demand = GetDemand();
}
```

sis3.c

Prof. Vittoria de Nitto Personè 15

15

necessità del troncamento: non ho limiti al numero di macchine comprate.
questo è recap troncamento

Discrete Simulation
Discrete RV applications

Truncation: examples

- In the previous example, no bound on number of autos purchased
- Can be made more realistic by <u>truncating</u> possible values
- Start with random variable X with possible values $\chi=\{0, 1, 2, \ldots\}$ and cdf $F(x)=\Pr(X \leq x)$
- want to restrict X to the finite range $0 \leq a \leq x \leq b < \infty$
- if $a > 0$, $\alpha = \Pr(X < a) = \Pr(X \leq a\text{-}1) = F(a\text{-}1)$
- $\beta = \Pr(X > b) = 1\text{-}\Pr(X \leq b) = 1\text{-}F(b)$
- $\Pr(a \leq X \leq b) = \Pr(X \leq b) - \Pr(X < a) = F(b) - F(a\text{-}1)$

  essentially, always true iff $F(b) \cong 1.0$ and $F(a\text{-}1) \cong 0.0$

Prof. Vittoria de Nitto Personè 16

16

nel modello 5, con istanze poisson e domande geometrica, faccio troncamento usando probabilità delle code. ho due possibilità: o parto dai valori, o da quanto voglio tagliare. qui uso un alfa piccolo e libreria rvms, da cui tiro fuori a e b, prendo inversa poisson per trovare valori corrispondenti ai bordi con massa alfa e beta.

Discrete Simulation
Discrete RV applications

The auto dealership

model 5

- the number of demand instances per time interval is *Poisson*(50)
- the demand per instance is *Geometric*(*p*) with p=0.375

For the *Poisson*(50) random variable I, determine a and b so that

$$\Pr(a \le I \le b) \cong 1.0$$

- use $\alpha = \beta = 10^{-6}$ and `rvms`

```
a = idfPoisson(50.0, α);
b = idfPoisson(50.0,1.0 - β);
```

- results: $a = 20$, $b = 87$
- consistent with the bounds produced by the conversion:

$\Pr(I < 20)$ = cdfPoisson(50.0, 19) $\cong 0.48 \times 10^{-6} < \alpha$
$\Pr(I > 87) = 1.0$ - cdfPoisson(50.0, 87) $\cong 0.75 \times 10^{-6} < \beta$

Prof. Vittoria de Nitto Personè 17

17

qui faccio una verifica usando una cumulativa per verificare che sia minore di alfa e di beta.

Discrete Simulation
Discrete RV applications

model 5

Effects of truncation

- truncating *Poisson*(50) to the range $\{20, \ldots, 87\}$ is insignificant: truncated and un-truncated random variables have (essentially) the same distribution

Prof. Vittoria de Nitto Personè 18

18

gli effetti del troncamento sono insignificanti (piccolo troncamento), quindi ho circa stessa distribuzione con code piu pesanti.

prendo Poisson (da libreria), genero domanda. quando d<20 assumo d=20, uguale per quaranta. cosi range è (20,40) quindi si accumula tutto su questi due valori, e non è corretto, sulle code voglio meno probabilità. 06/05/21

Troncamento deve essere fatto correttamente, e questo non è corretto.

model 3

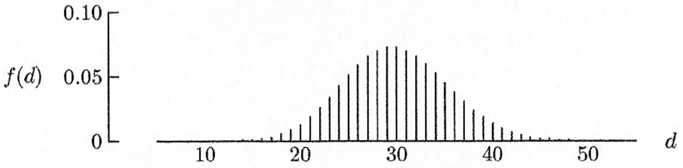
Discrete Simulation
Discrete RV applications

A *Poisson*(30) model

- *Binomial*(*n*,*p*) ≈ *Poisson*(*np*) for large n
- if *Binomial*(100,0.3) is realistic, should also consider *Poisson*(30)
- the function `GetDemand` in sis2 would be:

```
long GetDemand(void) {
        return (Poisson(30));
}
```

- μ=30, $\sigma = \sqrt{30} \cong 5.5$ and the pdf has slightly "heavier" tails:



Prof. Vittoria de Nitto Personè 19

19

ESEMPIO TRONCAMENTO ERRATO

model 3

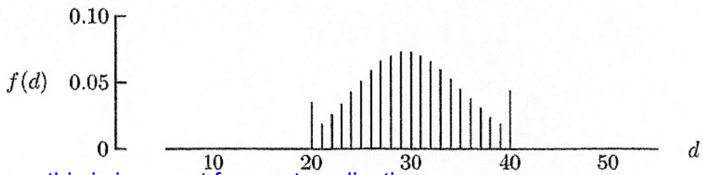
Discrete Simulation
Discrete RV applications

The auto dealership
incorrect truncation

- use *Poisson*(30) demand model in program sis2
- truncate the demand to the range $20 \leq d \leq 40$

```
d = Poisson(30.0);
if (d < 20)
        d = 20;
if (d > 40)
        d = 40;
return d;
```

- original left and right tails "grouped together" at 20 and 40



- this is incorrect for most applications

Prof. Vittoria de Nitto Personè 20

20

Vedendo Pareto e bounded pareto, sono rinormalizzazioni su un range.
presa la Poisson, devo calcolarmi la massa nel range che voglio scegliere.
qui range è (20,40). facendo tale somma ho massa probabilità 0.94
troncando la densità e normalizzandola nello spazio F(40) - F(19)

model 3

Discrete Simulation
Discrete RV applications

Truncation by cdf modification

- troncate *Poisson*(30) to range $20 \le d \le 40$
- the *Poisson*(30) pdf is:

$$f(d) = \exp(-30)\frac{30^d}{d!} \qquad d = 0,1,2,\ldots$$

$$\Pr(20 \le D \le 40) = F(40) - F(19) = \sum_{d=20}^{40} f(d) \cong 0.945817$$

- compute a new truncated random variable D_t with pdf $f_t(d)$

$$f_t(d) = \frac{f(d)}{F(40) - F(19)} \qquad d = 20,21,\ldots,40$$

Prof. Vittoria de Nitto Personè 21

21

qui riapplico le formule considerando il troncamento.
variabilità si è ridotta (meno valori possibili)

model 3

Discrete Simulation
Discrete RV applications

Truncation by cdf modification

- the corresponding truncated cdf is

$$F_t(d) = \sum_{t=20}^{d} f_t(t) = \frac{F(d) - F(19)}{F(40) - F(19)} \qquad d = 20,21,\ldots,40$$

- mean and standard deviation of D_t

$$\mu_t = \sum_{d=20}^{40} d f_t(d) \cong 29.841 \qquad \sigma_t = \sqrt{\sum_{d=20}^{40} (d - \mu_t)^2 f_t(d)} \cong 4.720$$

- mean and standard deviation of *Poisson*(30)

$$\mu = 30.0 \qquad \sigma = \sqrt{30} \cong 5.477$$

Prof. Vittoria de Nitto Personè 22

22

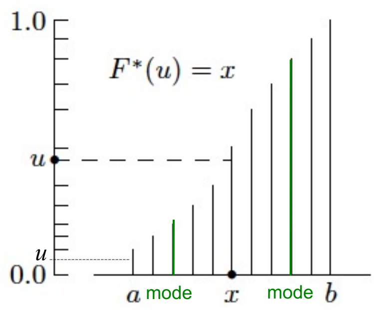
potremmo generare un u e fare troncamento come abbiamo già visto.

model 3

Truncation by cdf modification

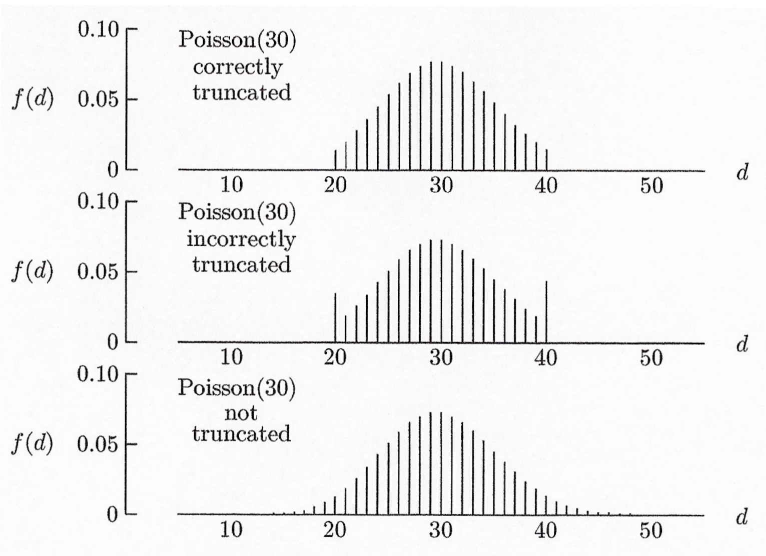
• A random variate truncated to $20 \le d \le 40$ can be generated by inversion, using the truncated cdf $F_t(\cdot)$ and the Algorithm 2

```
u = Random();
d = 30;
if (Ft(d) <= u)
      while (Ft(d) <= u)
            d++;
else if (Ft(20) <= u)
            while (Ft(d-1) > u)
                  d--;
      else
            d = 20;
return d;
```



23

Illustration of pdfs



24

il primo grafico è Poisson troncata correttamente, normalizzata ma con stessa forma (tolto prima di 20 e dopo 40).
Il secondo grafico ERRATO è accumulare sui punti 20 e 40.
il terzo grafico è non troncato.

Discrete Simulation
Discrete RV applications

Truncation - conclusion

see the book for the following

1. the truncation by cdf modification in general
2. a different approach called *truncation by constrained inversion*
3. the simple technique *truncation by acceptance-rejection*

Prof. Vittoria de Nitto Personè 25

25

chi fa questo mestiere, con queste valutazioni, deve essere familiare con:
che cosa fanno le distribuzioni (lancio dado, binomiale, etc...)

Important points

The modeler should be familiar with

- How these distribution arise
- The support, χ
- The mean, μ
- The variance, σ^2
- The shape of the pdf

- how these distributions relate to one another

Prof. Vittoria de Nitto Personè 26

26

Exercises

- Generating Discrete RV, use of libraries: exerc. 6.2.4
- Truncation: exerc. 6.3.1, 6.3.2