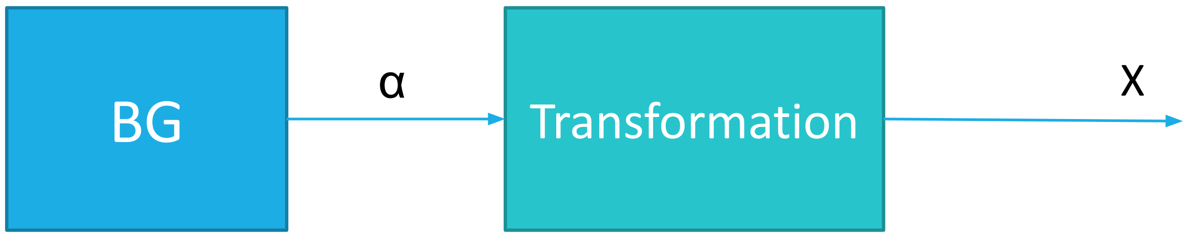
Simulation

STOCHASTIC MODELING

Random objects and base generator

1. Random events
2. Random variables
3. Random processes
4. Random point processes



Discrete RV

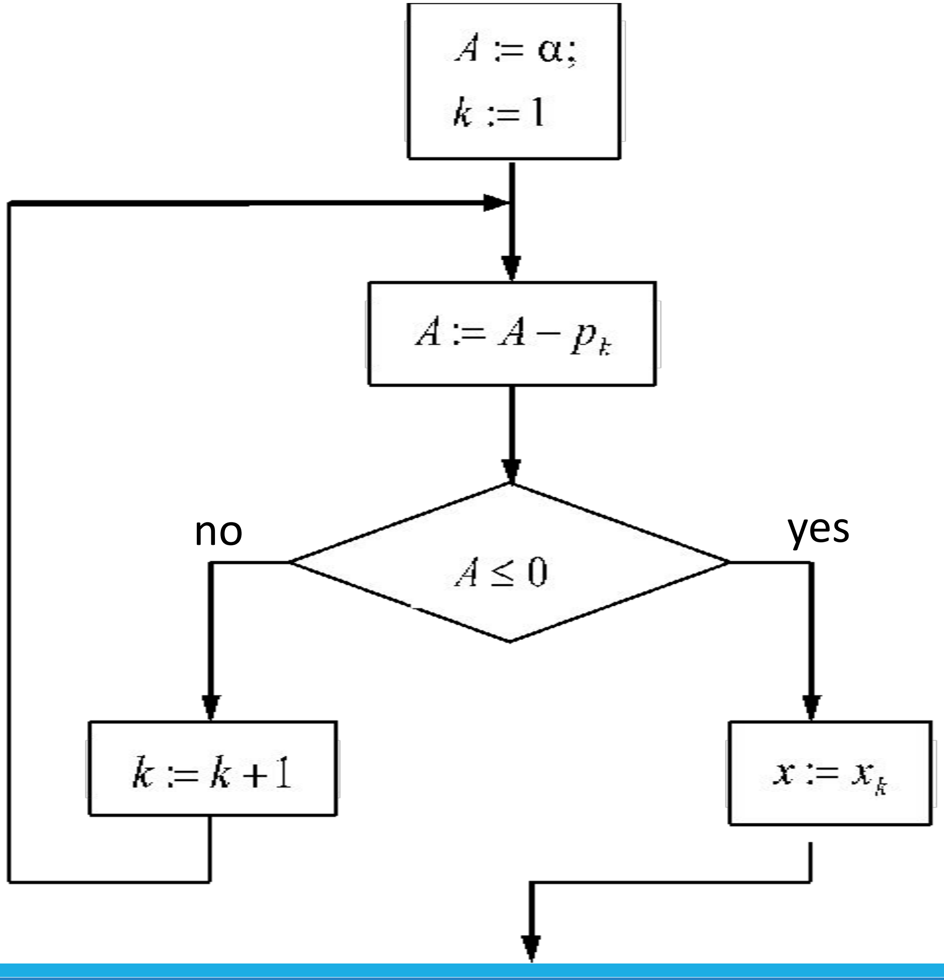
METHODS OF GENERATION

Discrete RV given by distribution

x	x_1	x_2	$\dots$	x_m
p	p_1	p_2	$\dots$	p_m

$$p_i = \mathrm{P}\{x = x_i\} \qquad \sum_{i=1}^{m} p_i = 1$$

$A_1 \quad A_2 \quad \dots \quad A_m$



Statistical processing for discrete RVs

x	x_1	x_2	$\dots$	x_m
p	p_1	p_2	$\dots$	p_m

Mathematical expectation (mean): $\mathrm{E}\,x = \sum_{i=1}^{m} p_i\, x_i$

Variance: $\mathrm{Var}\,x = \mathrm{D}\,x = \sum_{i=1}^{m} p_i (x_i - \mathrm{E}\,x)^2 = \sum_{i=1}^{m} p_i x_i^2 - (\mathrm{E}\,x)^2$

Relative frequencies

$$\hat{p}_i = \frac{n_i}{N}$$

n_i is the number of appearances of value *i* (they are named as frequencies),
N is the total number of trials

Empiric expectation (average):

$$\hat{\mathrm{E}}\,x = \sum_{i=1}^{m} \hat{p}_i\, x_i$$

Empiric variance:

$$\hat{\mathrm{D}}\,x = \sum_{i=1}^{m} \hat{p}_i x_i^2 - (\hat{\mathrm{E}}\,x)^2$$

Absolute errors:

$$\Delta_{\mathrm{E}} = \left|\hat{\mathrm{E}}\,x - \mathrm{E}\,x\right|$$

$$\Delta_{\mathrm{D}} = \left|\hat{\mathrm{D}}\,x - \mathrm{D}\,x\right|$$

Relative errors:

$$\delta_{\mathrm{E}} = \frac{\Delta_{\mathrm{E}}}{|\mathrm{E}\,x|} \qquad \delta_{\mathrm{D}} = \frac{\Delta_{\mathrm{D}}}{|\mathrm{D}\,x|}$$

Statistical processing for discrete RVs

Chi-squared test

$$\hat{X}_N^2 = \sum_{i=1}^{m} \frac{n_i^2}{Np_i} - N$$

Hypothesis that empiric distribution corresponds to the theoretical one is not true if and only if

$$\hat{X}_N^2 > \chi^2_{1-\alpha,m-1}$$

α is a significance level

Critical values of the Chi-square distribution with *d* degrees of freedom

	Probability of exceeding the critical value						
m	0.05	0.01	0.001	*m*	0.05	0.01	0.001
1	3.841	6.635	10.828	11	19.675	24.725	31.264
2	5.991	9.210	13.816	12	21.026	26.217	32.910
3	7.815	11.345	16.266	13	22.362	27.688	34.528
4	9.488	13.277	18.467	14	23.685	29.141	36.123
5	11.070	15.086	20.515	15	24.996	30.578	37.697
6	12.592	16.812	22.458	16	26.296	32.000	39.252
7	14.067	18.475	24.322	17	27.587	33.409	40.790
8	15.507	20.090	26.125	18	28.869	34.805	42.312
9	16.919	21.666	27.877	19	30.144	36.191	43.820
10	18.307	23.209	29.588	20	31.410	37.566	45.315

Laboratory #9

ASSIGNMENT:

Simulation of discrete random variable

- Implement an algorithm for conducting a series of experiments to simulate a discrete random variable specified by the distribution
- Calculate empirical probabilities, sample mean and variance, their relative errors
- Calculate the chi-squared statistic and apply the chi-squared test for different values of N (N = 10, 100, 1,000, 10,000)
- Draw a conclusion

Laboratory #9

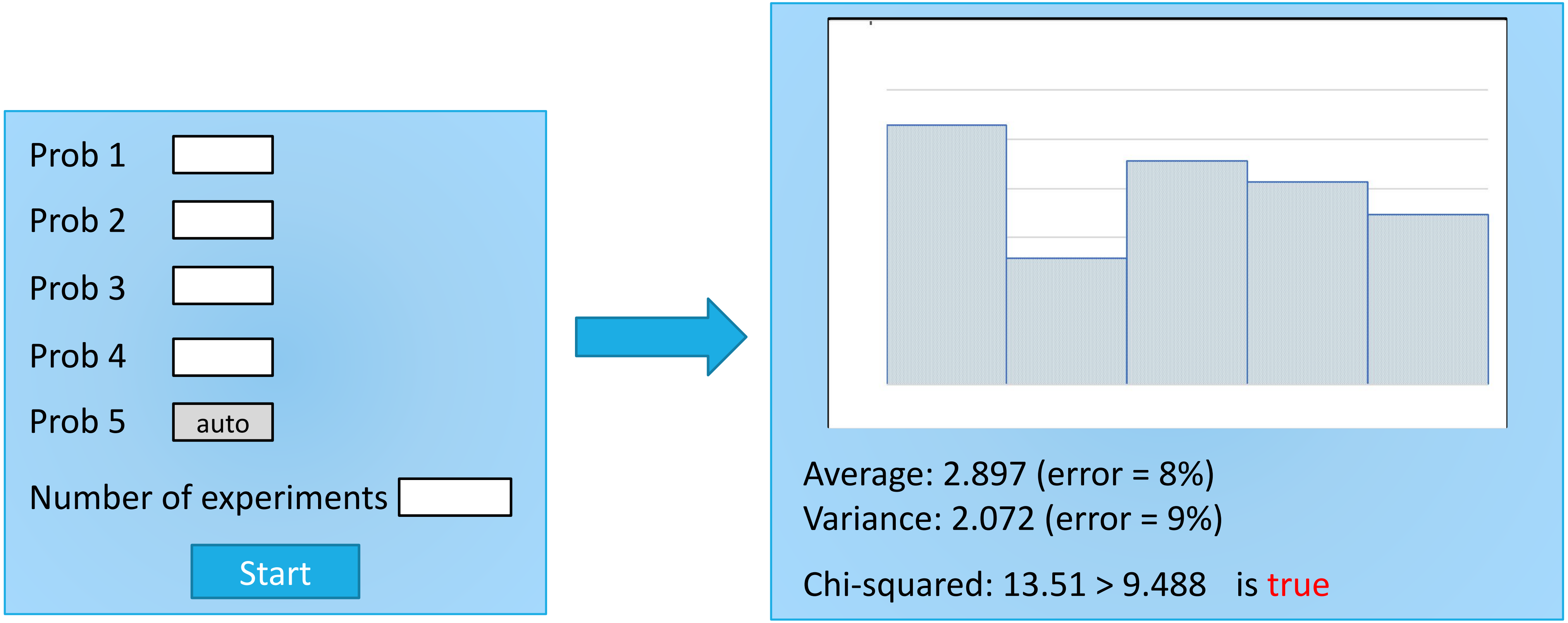
ЗАДАНИЕ:

Имитационное моделирование дискретных случайных величин

- Реализовать алгоритм проведения серии экспериментов по генерации дискретной случайной величины, заданной рядом распределения
- Вычислить эмпирические вероятности, выборочные среднее и дисперсию, их относительные погрешности
- Вычисление статистику хи-квадрат и применить критерий хи-квадрат при разных значениях N (N = 10, 100, 1 000, 10 000)
- Сделать вывод

Laboratory #9

UI PROTOTYPE



Uniform discrete distribution

FROM 0 TO n:

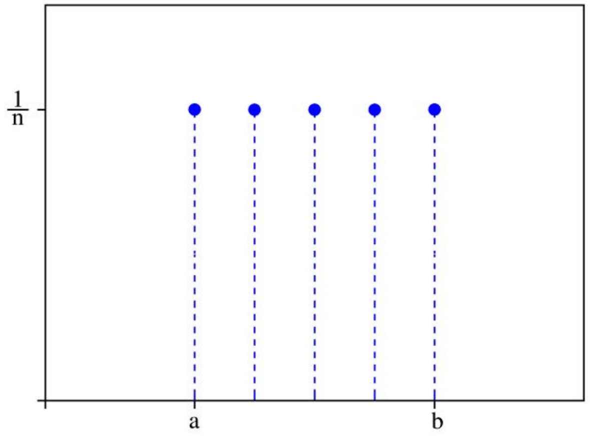
x	0	1	…	n
p	$\frac{1}{n+1}$	$\frac{1}{n+1}$	…	$\frac{1}{n+1}$

GENERATOR:

$$x = \mathrm{Int}(\alpha \cdot (n+1))$$

Int is a truncating operation

FROM a TO b:



1. Set $n = b - a$
2. Use $x = \mathrm{Int}(\alpha \cdot (n+1))$
3. Calculate $x = x + a$

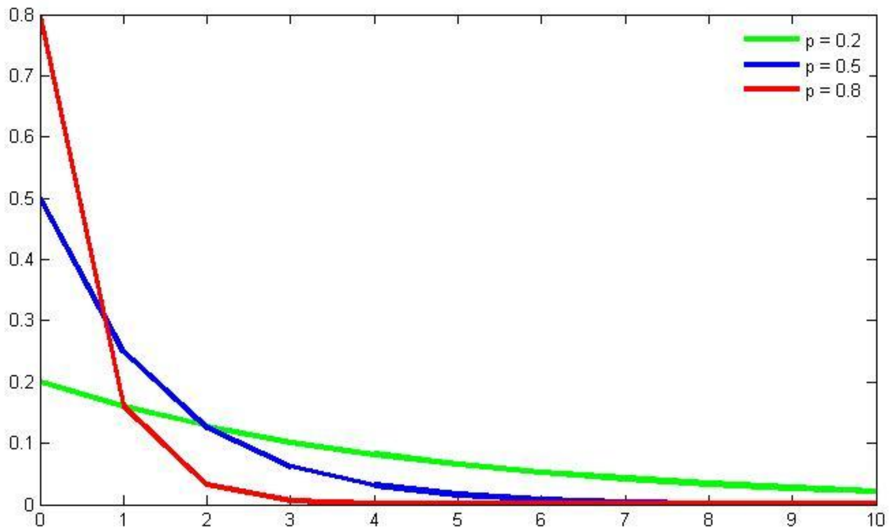
For example, if x from {1, 2, .., n} then use formula: $x = \mathrm{Int}(\alpha \cdot n) + 1$

Geometric distribution

The probability distribution of
the number of failures before the first success

$X \in \{0, 1, 2, 3, \ldots\}$

p is a probability of success in one trial



$$\mathrm{P}\{x = m\} = p(1 - p)^m$$

GENERATOR: $$x = \mathrm{Int}\left(\frac{\ln \alpha}{\ln(1 - p)}\right)$$

Negative binomial distribution

The probability distribution of
the number of failures before the *r*-th success

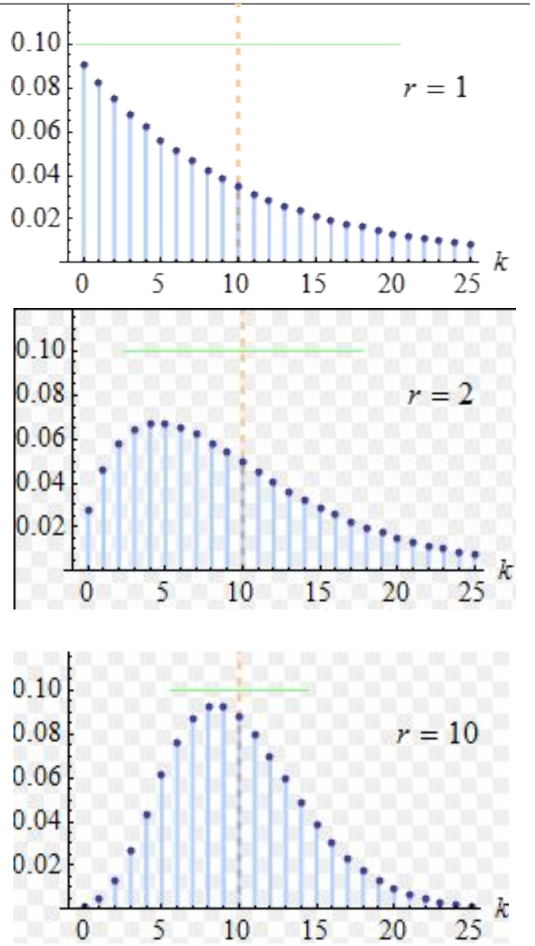
$X \in \{0, 1, 2, 3, \dots\}$

p is a probability of success in one trial

$$P\{x = m\} = C_{m+r-1}^{m} \cdot p^{r} (1-p)^{m}$$

GENERATOR:

$$x = \sum_{i=1}^{r} \mathrm{Int}\left(\frac{\ln \alpha_i}{\ln(1-p)} \right)$$



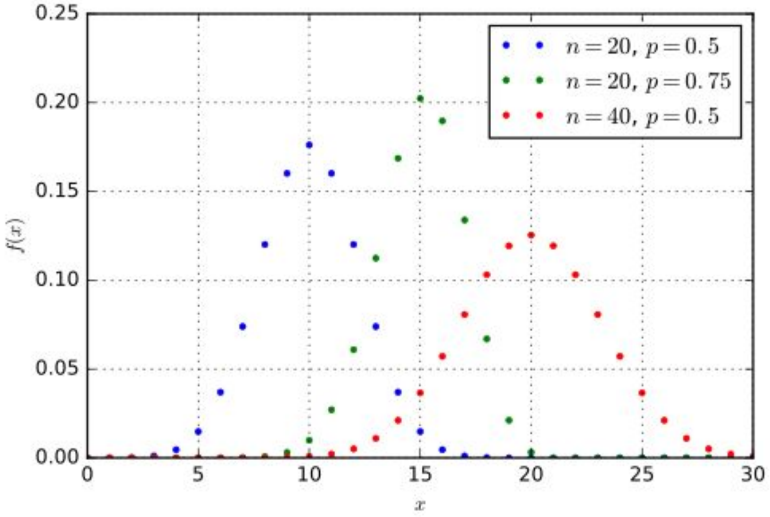
Binomial distribution

The probability distribution of
the number of successes in *n* trials

$X \in \{0, 1, 2, \ldots, n\}$

p is a probability of success in one trial

$$P\{x = m\} = C_n^m p^m (1-p)^{n-m}$$

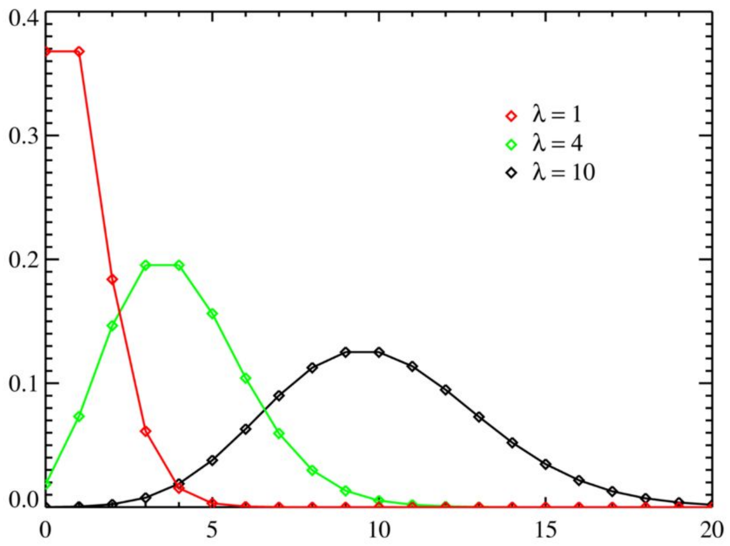


GENERATOR: $$x = \sum_{k=1}^{n} \theta(p - \alpha_k)$$ where $$\theta(x) = \begin{cases} 1, x \geq 0, \\ 0, x < 0. \end{cases}$$ is a Heaviside step function

Poisson distribution

The probability distribution of
the number of events occurred in a fixed interval of time if these events occur independently and with a known constant rate λ

$X \in \{0, 1, 2, 3, \dots\}$



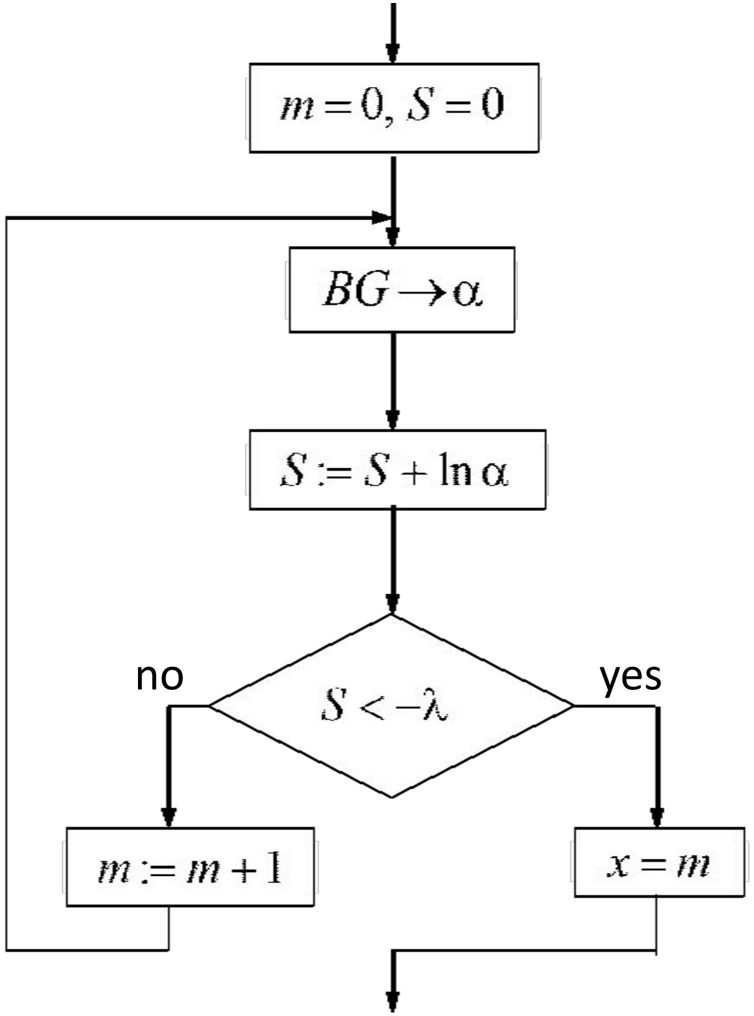
$$P\{x = m\} = \frac{\lambda^m}{m!} e^{-\lambda}$$

$$P\{x = m\} = \frac{(\lambda t)^m}{m!} e^{-\lambda t}$$

Poisson distribution generator

GENERATOR:

$$x = \min\left\{ m : \sum_{i=0}^{m} \ln \alpha_i < -\lambda \right\}$$



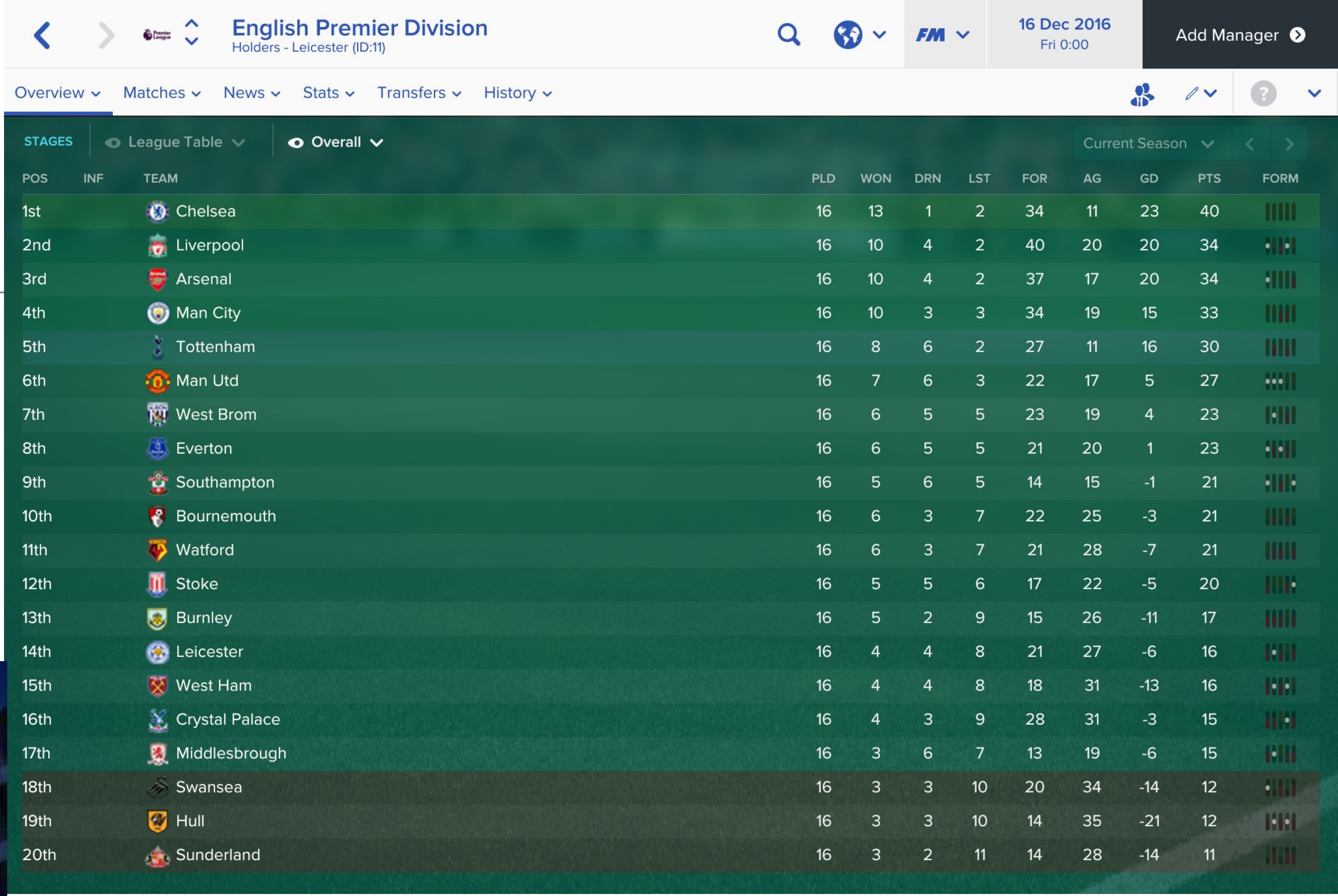
Laboratory #10

ASSIGNMENT:

Basketball

- ~~Football~~ manager game

Laboratory #10



POS	INF	TEAM	PLD	WON	DRN	LST	FOR	AG	GD	PTS	FORM
1st		Chelsea	16	13	1	2	34	11	23	40	
2nd		Liverpool	16	10	4	2	40	20	20	34	
3rd		Arsenal	16	10	4	2	37	17	20	34	
4th		Man City	16	10	3	3	34	19	15	33	
5th		Tottenham	16	8	6	2	27	11	16	30	
6th		Man Utd	16	7	6	3	22	17	5	27	
7th		West Brom	16	6	5	5	23	19	4	23	
8th		Everton	16	6	5	5	21	20	1	23	
9th		Southampton	16	5	6	5	14	15	-1	21	
10th		Bournemouth	16	6	3	7	22	25	-3	21	
11th		Watford	16	6	3	7	21	28	-7	21	
12th		Stoke	16	5	5	6	17	22	-5	20	
13th		Burnley	16	5	2	9	15	26	-11	17	
14th		Leicester	16	4	4	8	21	27	-6	16	
15th		West Ham	16	4	4	8	18	31	-13	16	
16th		Crystal Palace	16	4	3	9	28	31	-3	15	
17th		Middlesbrough	16	3	6	7	13	19	-6	15	
18th		Swansea	16	3	3	10	20	34	-14	12	
19th		Hull	16	3	3	10	14	35	-21	12	
20th		Sunderland	16	3	2	11	14	28	-14	11	

Use Poisson distribution for the number of goals in a match

Statistical processing for discrete RVs with infinite number of values

only for chi-squared criteria!

X: $$\sum_{i=X}^{\infty} p_i \le \min_{i<X} p_i$$

Example (geometric distribution p=0.8): $\mathrm{P}\{x = m\} = p(1-p)^m$

$$p_0 = 0.8; \qquad p_1 = 0.8 \text{ x } 0.2 = 0.16; \qquad p_2 = 0.8 \text{ x } 0.2 \text{ x } 0.2 = 0.032; \qquad \sum_{i=3}^{\infty} p_i = 0.008.$$

“Similar” distribution:

0	1	2	>2
0.8	0.16	0.032	0.008