Application Examples of Probabilities, Random Variables and Markov Chains

Desempenho e Dimensionamento de Redes
Universidade de Aveiro

Jorge Catarino, Ricardo Azevedo

universidade de aveiro
theoria poiesis praxis

Version 1.0

Application Examples of Probabilities, Random Variables and Markov Chains

DETI
Desempenho e Dimensionamento de Redes
Universidade de Aveiro

(85028) jorge.catarino@ua.pt

(84730) azevedoricardo@ua.pt

10 de Abril, 2021

Conteúdo

Lista de Figuras

Capítulo 1

Tarefa 4

Neste capítulo iremos em primeiro lugar expor os resultados da Tarefa 4 e o raciocínio para os alcançar. Em seguida, será apresentado o código desenvolvido e, quando considerado necessário, uma reflexão sobre os resultados obtidos.

1.1 Alínea a)

1.1.1 Método e Resultados

Para determinar as probabilidades da ligação estar em estado normal e no estado de interferência, começamos por calcular as probabilidades de cada estado da cadeia de Markov. Sabendo que este é um processo de nascimento e morte, estas podem ser calculadas a partir de:

$$\pi_n = \frac{(\lambda_0 \lambda_1 ... \lambda_n - 1)}{(\mu_1 \mu_2 ... \mu_n)} \lambda_0, n >= 1$$

Sabendo então as probabilidades associadas a cada estado do processo, o passo seguinte é somar os valores associados aos estados em que a ligação está em estado normal e os que a ligação se encontra em estado de interferência. Feito isto, obtemos então os seguintes resultados:

$$P(Normal) = 0.99998;$$

$$P(Interference) = 1.5784 \times 10^{-5};$$

1.1.2 Código

```
%4a
% Interference >=10^-3
% Normal < 10^-3

transitions = [8/600,5/200,2/50,1/5];

% calculo da probabilidade limite de processos de nascimento e morte

 tmp = transitions(1);
for i = 2:4
    sum = transitions(i);
    for j = 1:(i-1)
        sum = sum * transitions(j);
    end
    tmp = tmp + sum;
end

res_1 = 1/(1+ tmp);
res_2 = res_1*transitions(1);
res_3 = res_2*transitions(2);
res_4 = res_3*transitions(3);
res_5 = res_4*transitions(4);

X = [res_1 res_2 res_3 res_4 res_5];

prob_normal = X(1) + X(2) + X(3);
prob_interf = X(4) + X(5);

disp(' ');
disp("Ex4.a.");
disp("Probability of the link being in normal state = " + prob_normal);
disp("Probability of the link being in interference state = " + prob_interf);
```

1.2 Alínea b)

1.2.1 Método e Resultados

Para determinar a média de bit error rate (BER) no estado normal e no estado de inferência, começamos por multiplicar a probabilidade de cada estado da cadeia de Markov pelos seus respectivos BER, o resultado é a média de BER de cada estado. Este resultado é então dado pela formula :

$$AverageBER(State) = BER(State) * P(State);$$

Sabendo então a média de BER de cada estado podemos então obter a média de BER de quando a ligação está em estado normal ou em estado de inferência somando as médias de BER de cada estado envolvente no estado normal ou de interferência e dividi-la pela probabilidade de a ligação estar neste estado. Com isto, obtemos então os seguintes resultados :

$$AverageBER(Normal) = \frac{AverageBER(10^{-6}) + AverageBER(10^{-5}) + AverageBER(10^{-4})}{P(Normal)} = 1.1509 \times 10^{-6};$$

$$AverageBER(Interference) = \frac{AverageBER(10^{-3}) + AverageBER(10^{-2})}{P(Interference)} = 0.0025;$$

1.2.2 Código

```
%4b
disp(' ');
disp("Ex4.b.");
average_normal = ((X(1) * 10^-6) + (X(2) * 10^-5) + (X
    (3) * 10^-4))/prob_normal;
average_interf = ((X(4)*10^-3) + (X(5)*10^-2))/
    prob_interf;
disp("Average ber of the link in normal state = " +
    average_normal);
disp("Average ber of the link in interference state =
    " + average_interf);
```

1.3 Alínea c)

1.3.1 Método e Resultados

A probabilidade da estação estar em modo normal dado que recebeu um pacote de dados com erros pode ser calculada a partir do Teorema de Bayes:

$$P(Normal|Errors) = \frac{P(Errors|Normal)P(Normal)}{P(Errors)} =$$
$$\frac{P(Errors|State_1)P(State_1) + P(Errors|State_2)P(State_2) + P(Errors|State_3)P(State_3)}{P(Errors)}$$

A probabilidade da ligação estar em estado normal, $P(Normal)$, já foi calculada num exercício anterior (Secção 1.1), faltando então só calcular as probabilidades do pacote ter erros sabendo que a ligação se encontra normal, $P(Errors|State_n)$, sendo $n <= 3$, e a probabilidade do pacote ter erros, $P(Errors)$.

Assumindo que os erros nos diferentes bits do pacote de dados são estatisticamente independentes, podemos afirmar que o número de erros nestes serão uma varíavel aleatória com distribuição binomial, sendo que probabilidade de sucesso corresponde ao BER e o número de experiência será então o número de bits do pacote.

Com isto, usando o BER correspondente a cada estado, o tamanho de um pacote e tendo $i = 0$, podemos obter $P(Errors|State_n)$ com a expressão:

$$P(Errors|State_n) = 1 - (1 - state\_n\_ber)^{packet\_size}$$

Tendo isto, resta calcular $P(Errors)$. Sabendo que a união dos acontecimentos que compõe o estado Normal ($10^{-6}, 10^{-5}, 10^{-4}$) e dos acontecimentos pertencentes ao estado de Interferência ($10^{-3}, 10^{-2}$) resulta no espaço de resultados S e que estes são mutuamente exclusivos, esta probabilidade pode então ser calculada com:

$$P(Errors) = \sum_{n=1}^{5} P(Errors|State_n)P(State_n)$$

Seguindo este método, podemos então calcular a probabilidade da ligação estar em estado normal em função do tamanho dos pacotes. O intervalo utilizado foi de 64 bytes a 200 bytes. O gráfico resultante da aplicação pode ser consultado na Figura 1.1.

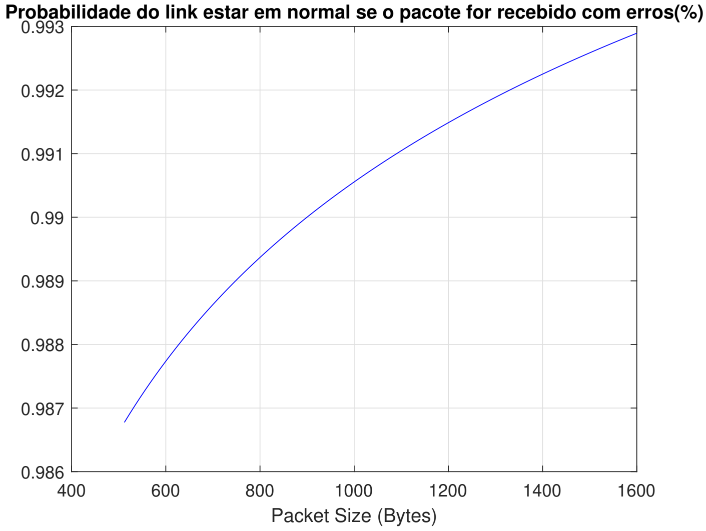

Figura 1.1: Resultado da alínea c)

1.3.2 Código

```
%4c

disp(' ');
disp("Ex4.c.");
disp("Check figure 1.");
% Packet size from 64 bytes to 200 bytes
n = linspace(64, 200) * 8;

ber_state = [(10^-6) (10^-5) (10^-4) (10^-3) (10^-2)];

%With Errors
prob_errors_in_state_1 = (1 - ((1 - ber_state(1)).^(n)));
prob_errors_in_state_2 = (1 - ((1 - ber_state(2)).^(n)));
prob_errors_in_state_3 = (1 - ((1 - ber_state(3)).^(n)));
prob_errors_in_state_4 = (1 - ((1 - ber_state(4)).^(n)));
prob_errors_in_state_5 = (1 - ((1 - ber_state(5)).^(n)));

% P(CondN|Errors) = P(Errors|N)P(N)/P(Errors)
%prob_normal_with_errors = (prob_errors_in_normal .* prob_normal)./prob_errors;
prob_errors_state_sum = (prob_errors_in_state_1*X(1)) +...
    (prob_errors_in_state_2*X(2)) +...
    (prob_errors_in_state_3*X(3)) +...
    (prob_errors_in_state_4*X(4)) +...
    (prob_errors_in_state_5*X(5));

prob_normal_with_errors = ((prob_errors_in_state_1*X(1)) + ...
    (prob_errors_in_state_2*X(2)) + (prob_errors_in_state_3*X(3)) ) ./ prob_errors_state_sum;

figure(1);
plot(n, prob_normal_with_errors, 'b-');
grid on;
title("Probabilidade do link estar em normal se o pacote for recebido com erros(%)");
xlabel('Packet Size (Bytes)');
print -depsc alinea4c
```

1.3.3 Discussão

Com base nos resultados, podemos concluir que a probabilidade da ligação estar em estado normal quando recebe um pacote com erros, em função do tamanho dos pacotes, é bastante elevada, tendendo para 1. Apesar de parecer contraditório à primeira vista, isto acontece porque a probabilidade da ligação estar em estado normal é também próxima de 1, o que faz com que a chance desta estar em modo de interferência tenda para 0, consequentemente reduzindo a probabilidade da ligação estar neste último modo quando recebe um pacote com erros.

1.4 Alínea d)

1.4.1 Método e Resultados

A probabilidade da estação estar em modo de interferência dado que recebeu um pacote de dados sem erros pode ser calculada a partir do Teorema de Bayes, de forma semelhante ao que foi feito na Secção 1.3:

$$P(Interference|No\_Errors) = \frac{P(No\_Errors|State_4)P(State_4) + P(No\_Errors|State_5)P(State_5)}{P(No\_Errors)}$$

Estes valores podem ser obtidos a partir dos que foram calculados, dado que:

$$P(No\_Errors|State_n) = 1 - P(Errors|State_n)$$
$$P(No\_Errors) = 1 - P(Errors)$$

Tendo este conhecimento, é então possível proceder ao cálculo das probabilidades pedidas. O gráfico resultante da sua aplicação do intervalo de 64 a 200 bytes pode ser consultado na Figura 1.2.

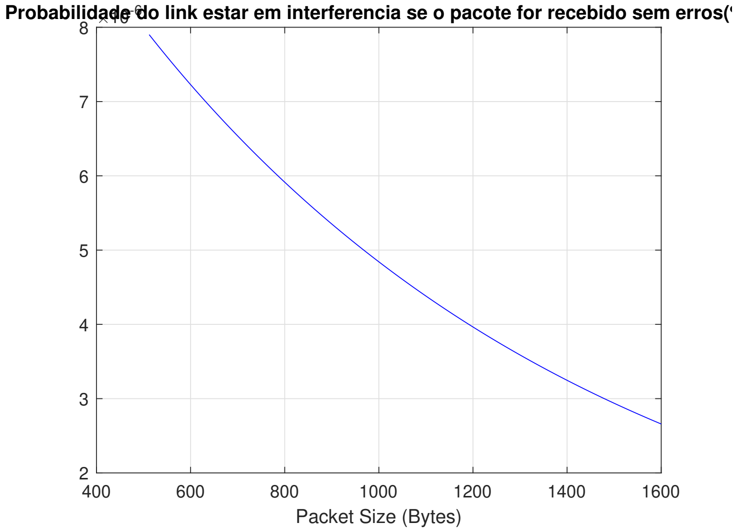

Figura 1.2: Resultado da alínea d)

1.4.2 Código

```
disp(' ');
disp("Ex4.d.");
disp("Check figure 2.");
% Packet size from 64 bytes to 200 bytes
n = linspace(64, 200) * 8;

% Without Errors
prob_no_errors_in_state_4 = 1-prob_errors_in_state_4;
prob_no_errors_in_state_5 = 1-prob_errors_in_state_5;

prob_no_errors_state_sum = 1 - prob_errors_state_sum;

% P(CondI|No_Errors) = P(No_Errors|I)P(I)/P(No_Errors)
prob_interf_no_errors = ((prob_no_errors_in_state_4*X
    (4)) + (prob_no_errors_in_state_5*X(5))) ./
    prob_no_errors_state_sum;

figure(2);
plot(n, prob_interf_no_errors, 'b-');
grid on;
title("Probabilidade do link estar em interferencia se
     o pacote for recebido sem erros(%)");
xlabel('Packet Size (Bytes)');
print -depsc alinea4d

save("task4variables.mat");
```

1.4.3 Discussão

Com base nos resultado, pode-se afirmar que a probabilidade da ligação estar em modo de interferência quando é recebido um pacote sem erros tende para 0, à medida que o tamanho do pacote aumenta. Há dois fatores que influenciam estes resultados, a baixa probabilidade da estação estar em modo de interferência e o BER neste modo.

A taxa de erro, o BER, em interferência é bastante elevado, o que significa que este modo tem uma têndencia para erro, diminuindo desta forma a probabilidade de receber um pacote sem erros.

Capítulo 2

Tarefa 5

Neste capítulo iremos em primeiro lugar expor os resultados da Tarefa 5 e o raciocínio para os alcançar. Em seguida, será apresentado o código desenvolvido e, quando considerado necessário, uma reflexão sobre os resultados obtidos.

2.1 Alínea a) e b)

2.1.1 Método e Resultados

Alínea a)

Um falso-positivo acontece quando a estação decide de forma errónea que a ligação está em estado de interferência, como por exemplo, quando recebe n pacotes de controlo com erro e esta está em estado normal. A probabilidade de acontecer um falso-positivo é a probabilidade de estar em estado normal sabendo que todas os pacotes de controlo foram recebidos com erro, que pode ser calculada com o Teorema de Bayes, de modo semelhante ao que já foi feito em exercícios anteriores:

$$P(False\_Positive) = \frac{P(Normal)P(Errors|Normal)}{P(Normal)P(Errors|Normal) + P(Interf)P(Errors|Interf)}$$

Como a chance de um pacote de controlo ter erros é um evento independente, a probabilidade de eventos sucessivos pode ser calculada como $(P(Errors|Normal))^n$, onde n é o número de pacotes consecutivos recebidos.

Com isto, é possível afirmar que a probabilidade calculada na Secção 1.3 é um caso particular desta alínea, com $n = 1$. Aplicando estes conhecimento, com n pertencente a um intervalo de 2 a 5, obtem-se então o gráfico representado na figura 2.1.

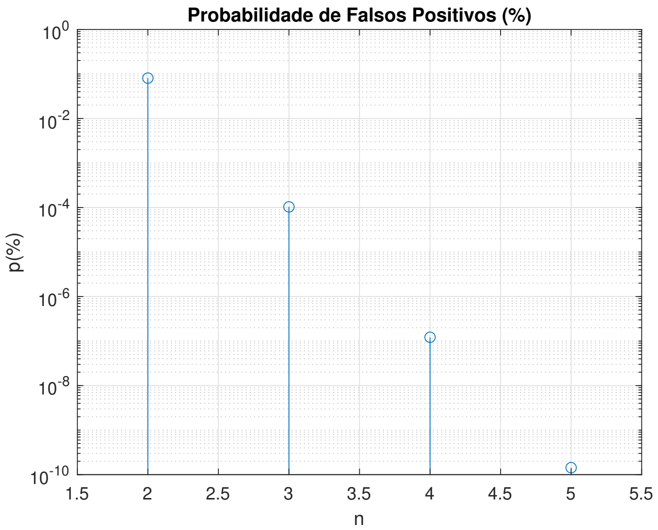

Figura 2.1: Resultado da alínea a)

Alínea b)

Um falso-negativo acontece quando a estação decide de forma errónea que a ligação está em estado de normal, como por exemplo, quando recebe n pacotes de controlo sem erro e esta está em modo de interferência. A probabilidade deste evento acontecer corresponde à probabilidade de estar em estado de interferência sendo que pelo menos um dos pacotes de controlo é recebido sem erros. De forma semelhante ao que foi feito na alínea anterior, esta pode calculada com o Teorema de Bayes:

$$P(False\_Negative) = \frac{P(Interf)P(No\_Errors|Interf)}{P(Normal)P(No\_Errors|Normal) + P(Interf)P(No\_Errors|Interf)}$$

Os valores de $P(No\_Errors|Interf)$ e $P(Normal)P(No\_Errors|Normal)+ P(Interf)P(No\_Errors|Interf)$ correspondem à negação de valores calculados na alínea anterior.

Aplicando estes conhecimentos, com n pertecente ao intervalo de 2 a 5, obtemos então o resultado figurado na Figura 2.2

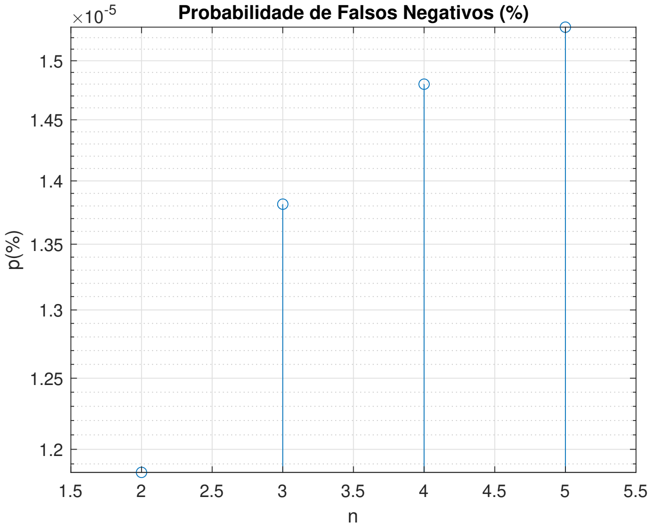

Figura 2.2: Resultado da alínea b)

2.1.2 Código

```
%Task5

% loading saved values from the last task - Run last task once to create
% the file
load("task4variables.mat");

%5a and b
% Each event is independent, so the probability of having errors for n
% frames can be given by (prob_no_errors_in_x)^n_frame
% Frame size is 64

prob_false_positive_n = zeros(1, 4);
prob_false_negative_n = zeros(1, 4);

% P(Errors|Normal) = P(Normal|Errors)P(Errors)/P(Normal);
% P(Errors|Interf) = P(Interf|Errors)P(Errors)/P(Normal);
prob_errors_in_normal = (prob_normal_with_errors.*prob_errors_state_sum)./prob_normal;
```

```matlab
prob_interf_with_errors = ((prob_errors_in_state_4*X(4)) + ...
    (prob_errors_in_state_5*X(5))) ./ prob_errors_state_sum;
prob_errors_in_interf = (prob_interf_with_errors.*prob_errors_state_sum)./prob_interf;

for n_frame = 2:5
    % Using variables from last exercise, p_errors_in_normal(1) corresponds
    % to the probability of the packet having errors in normal state for a
    % packet size of 64
    prob_errors_in_normal_n = prob_errors_in_normal(1) ^ n_frame;
    prob_errors_in_interf_n = prob_errors_in_interf(1) ^ n_frame;

    prob_no_errors_in_interf_n = 1 - prob_errors_in_interf_n;

    % prob_falsep = P(Normal)P(Errors|Normal)/P(Normal)P(Errors|Normal) +
    % P(Interf)P(Errors|Interf)

    % prob_falsen =
    % P(Interf)P(No_Errors|Interf)/P(Normal)P(No_Errors|Normal) +
    % P(Interf)P(No_Errors|Interf)

    c_index = n_frame-1;

    prob_false_positive_n(c_index) = (prob_normal * prob_errors_in_normal_n)/...
        ((prob_normal * prob_errors_in_normal_n) +...
        (prob_interf * prob_errors_in_interf_n));

    prob_false_negative_n(c_index) = (prob_no_errors_in_interf_n) * prob_interf/...
        (1 - ((prob_normal * prob_errors_in_normal_n) +...
        (prob_interf * prob_errors_in_interf_n)));

    disp("n = " + n_frame + "; Probability of False Positive = " + ...
         prob_false_positive_n(c_index));
end

disp("Ex5.a.");
disp("Check figure 1.");
```

```
figure(1);
stem([2,3,4,5], prob_false_positive_n);
set(gca,'yscal','log');
title("Probability of False Positives (%)");
xlabel('n');
ylabel('p(%)');
xlim([1.5, 5.5]);
grid on;
print -depsc alinea5a

disp("Ex5.b.");
disp("Check figure 2.");
figure(2);
stem([2,3,4,5], prob_false_negative_n);
set(gca, 'yscal', 'log');
title("Probability of False Negatives (%)");
xlabel('n');
ylabel('p(%)');
xlim([1.5, 5.5]);
grid on;
print -depsc alinea5b
```

2.1.3 Discussão

Com base nos resultados obtidos em ambas as alíneas e tendo em conta os resultados anteriores relativos à probabilidade da ligação estar em estado normal se o pacote for recebido com erros, é fácil perceber a razão de a probabilidade de um falso-positivo ocorrer com poucos pacotes. Isto deve-se ao facto da probabilidade da ligação estar em estado normal recebendo um pacote com erros tender para 100%.

À medida que se vai aumentando o número de pacotes de controlo, a perceção das estações em relação ao estado em que a ligação se encontra é melhorada, comprovada pela diminuição da probabilidade de existência de erros estando a ligação em estado normal. Isto faz com que a probabilidade de falsos-positivos diminua drasticamente, dado que as estações conseguem perceber que a ligação está no estado normal mesmo recebendo erros.

No caso dos falsos-negativos, esta probabilidade está relacionada com a probabilidade da ligação estar no estado de interferência não tendo recebendo erros. Visto que esta probabilidade é extremamente baixa (como foi confirmado nos exercícios anteriores) é então normal que a probabilidade de falsos-negativos seja também baixa.

Ao contrário do que acontece com a probabilidade de falsos-positivos, o incremento do número de pacotes de controlo pouco afecta a probabilidade de falsos-negativos.

2.2 Alínea c)

2.2.1 Discussão

Analisando os resultados da alínea a) e b) do exercício 5 e tendo em conta que as probabilidades de falsos-positivos e falsos-negativos tem igual importância, é possível concluir que o valor de n ideal para o sistema em questão é $n = 5$.

Apesar da probabilidade de falsos-negativos atingir um máximo em $n = 5$, no intervalo $[2, 5]$, esta ainda é relativamente baixa. De facto, a diferença entre o valor máximo e mínimo $p(n)$ deste intervalo é apenas, aproximadamente 0.32×10^{-5}. O que acontece é que a diferença nas probabilidades em função de n acaba por ser desprezável, não afectando consideravelmente a precisão do sistema qualquer que seja o n escolhido.

Já a probabilidade de falsos-negativos atinge um mínimo em $n = 5$, no intervalo $[2, 5]$, que é aproximadamente $562, 731, 902.38$ vezes menor que o valor máximo desta probabilidade, neste intervalo. As probabilidades desta função acabam por variar excessivamente quando se varia o n. Desta forma, esta acaba por ser o factor decisivo na escolha do n ideal para este sistema.