MATD49-Estatística não paramétrica

Testes de Aderência

Testes de aderência - Adequação do ajuste

Os testes de aderência (*goodness-of-fit*) avaliam se a distância da distribuição dos dados observados é significativa em relação a uma distribuição de referência. Veremos nesta aula os seguintes testes de aderência:

- Teste de Kolmogorov-Smirnov;
- Teste de Lilliefors;
- Teste de Shapiro-Wilk;
- Teste de Anderson-Darling.

Teste de Kolmogorov-Smirnov

- Avalia se os dados amostrais se aproximam razoavelmente de uma determinada distribuição.
- O teste de Kolmogorov-Smirnov se baseia na distância máxima entre a distribuição observada e a distribuição teórica de referência.

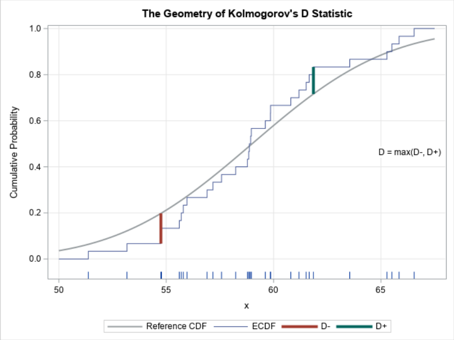


Figura 1: Diferença máxima entre a distribuição teórica e observada. Fonte:blogs.sas.com

Pressupostos

- O nível de mensuração da variável deve seguir ao menos uma escala ordinal (mais comum em intervalar ou de razão).
- A distribuição a ser testada deve ser plenamente especificada.
- Aplicável, sem restrições, a pequenas amostras.
- A amostra é aleatória.

Hipóteses: Caso Bilateral

Seja $F^*(x)$ a distribuição acumulada esperada (teórica) e $F_o(x)$ a distribuição acumulada observada. Queremos testar se a amostra é proveniente de uma população com distribuição plenamente especificada $F(x)$, ou seja:

Teste Bilateral:

H_0: $F_o(x) = F^*(x)$ para todo $-\infty < x < \infty$.

H_1: $F_o(x) \neq F^*(x)$ para pelo menos um valor de x.

Estatística de Teste I

- Seja $S(x)$ a distribuição de frequências acumuladas construídas a partir da amostra ordenada de tamanho n.
- $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ denotam as estatísticas de ordem da amostra aleatória.
- A função de distribuição empírica é definida como:

$$S(x) = \begin{cases} 0, & \text{se } x < X_{(1)} \\ \dfrac{k}{n}, & \text{se } X_{(k)} \leq x < X_{(k+1)} \text{ para } k = 1, 2, \ldots, n-1 \\ 1, & \text{se } x \geq X_{(n)} \end{cases} .$$

Obs.: k pode ser visto como o número de observações não superiores ao valor de x.

Estatística de Teste II

- A variável aleatória $S(X)$, que é a função de distribuição empírica de uma amostra aleatória $X_1, X_2, \ldots, X_n$ de uma distribuição $F^*(x)$, tem distribuição de probabilidade definida por:

$$P\left(S(X) = \frac{k}{n}\right) = \binom{n}{k}(F^*(x))^k(1 - F^*(x))^{n-k}, \; k = 0, 1, \ldots, n.$$

- $S(X)$ converge em probabilidade para $F^*(x) \Rightarrow S(X)$ é um estimador consistente de $F^*(x)$.
- A estatística do teste é dada por:

$$D = \sup_x \left| F^*(x) - S(x) \right|,$$

corresponde a maior distância vertical entre $F^*(x)$ e $S(x)$, é chamada *estatística de Kolmogorov-Smirnov* para uma amostra.

- Compara-se o valor observado D com o valor crítico tabelado.

Estatística de Teste III

TABLE A13 Quantiles of the Kolmogorov Test Statistic[a]

One-Sided Test											
	$p = 0.90$	0.95	0.975	0.99	0.995		$p = 0.90$	0.95	0.975	0.99	0.995
Two-Sided Test											
	$p = 0.80$	0.90	0.95	0.98	0.99		$p = 0.80$	0.90	0.95	0.98	0.99
$n = 1$	0.900	0.950	0.975	0.990	0.995	$n = 21$	0.226	0.259	0.287	0.321	0.344
2	0.684	0.776	0.842	0.900	0.929	22	0.221	0.253	0.281	0.314	0.337
3	0.565	0.636	0.708	0.785	0.829	23	0.216	0.247	0.275	0.307	0.330
4	0.493	0.565	0.624	0.689	0.734	24	0.212	0.242	0.269	0.301	0.323
5	0.447	0.509	0.563	0.627	0.669	25	0.208	0.238	0.264	0.295	0.317
6	0.410	0.468	0.519	0.577	0.617	26	0.204	0.233	0.259	0.290	0.311
7	0.381	0.436	0.483	0.538	0.576	27	0.200	0.229	0.254	0.284	0.305
8	0.358	0.410	0.454	0.507	0.542	28	0.197	0.225	0.250	0.279	0.300
9	0.339	0.387	0.430	0.480	0.513	29	0.193	0.221	0.246	0.275	0.295
10	0.323	0.369	0.409	0.457	0.489	30	0.190	0.218	0.242	0.270	0.290
11	0.308	0.352	0.391	0.437	0.468	31	0.187	0.214	0.238	0.266	0.285
12	0.296	0.338	0.375	0.419	0.449	32	0.184	0.211	0.234	0.262	0.281
13	0.285	0.325	0.361	0.404	0.432	33	0.182	0.208	0.231	0.258	0.277
14	0.275	0.314	0.349	0.390	0.418	34	0.179	0.205	0.227	0.254	0.273
15	0.266	0.304	0.338	0.377	0.404	35	0.177	0.202	0.224	0.251	0.269
16	0.258	0.295	0.327	0.366	0.392	36	0.174	0.199	0.221	0.247	0.265
17	0.250	0.286	0.318	0.355	0.381	37	0.172	0.196	0.218	0.244	0.262
18	0.244	0.279	0.309	0.346	0.371	38	0.170	0.194	0.215	0.241	0.258
19	0.237	0.271	0.301	0.337	0.361	39	0.168	0.191	0.213	0.238	0.255
20	0.232	0.265	0.294	0.329	0.352	40	0.165	0.189	0.210	0.235	0.252
						Approximation for $n > 40$	$\frac{1.07}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.52}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

Figura 2: Tabela dos valores críticos para o teste de Kolmogorov-Smirnov. Fonte: [Conover, 1996].

Estatística de Teste IV

Decisão do Teste:

Para um nível de significância α, rejeitar H_0 se a estatística D exceder ao valor do quantil de $1-\alpha$ da Tabela A13 na Figura 2.

Estatística de Teste V

Observação:

Como a função de distribuição empírica $S(x)$ é descontínua e a função de distribuição hipotética é contínua, uma alternativa é considerar:

$$\begin{aligned} D_1 &= \sup_x |F^*(x_{(i)}) - S(x_{(i)})| \\ D_2 &= \sup_x |F^*(x_{(i)}) - S(x_{(i-1)})| \end{aligned}$$

Essas estatísticas medem as distâncias (vertical) entre os gráficos das duas funções, teórica e empírica, nos pontos $x_{(i-1)}$ e $x_{(i)}$. Com isso, podemos utilizar como estatística de teste:

$$D = \max(D_1, D_2).$$

Aspecto Computacional

No R:

```
ks.test(x, y, alternative = c("two.sided", "less", "greater"),...)
# x: Dados observados.
# y: Quantis teoricos ou nome da dist. (apenas continuas, e.g.: pnorm).
# alternative: tipo de teste que se deseja realizar
# ...: parametros adicionais da distribuicao definida em y.
```

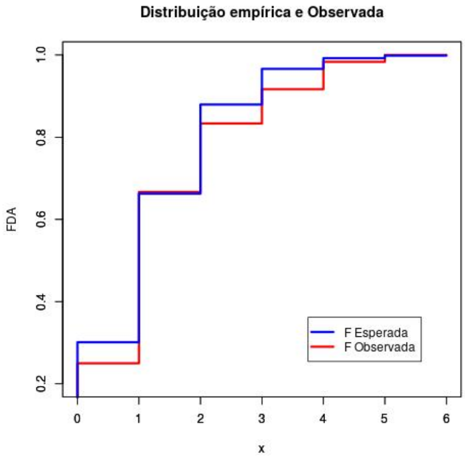
No SAS:

```
proc npar1way;
run;
```

Exemplo

Verifique se os dados abaixo podem ser ajustados por uma distribuição de Poisson com média igual a 1.2. Considere $\alpha = 0.05$.

X_i	f_i	f_i^*	F_i	F_i^*	$\|F_i^* - F_i\|$	$\|F_i^* - F_{i-1}\|$
0	15	18	0.25	0.3	0.052119	-
1	25	22	0.67	0.66	0.00404	0.41263
2	10	13	0.83	0.88	0.04615	0.21282
3	5	5	0.92	0.97	0.04956	0.1329
4	4	2	0.98	0.99	0.00892	0.07559
5	1	0	1	1	0.0015	0.01517



«Código R»

Exemplo

Verifique se os dados abaixo podem ser ajustados por uma distribuição de Poisson com média igual a 1.2. Considere $\alpha = 0.05$.

X_i	f_i	f_i^*	F_i	F_i^*	$\lvert F_i^* - F_i \rvert$	$\lvert F_i^* - F_{i-1} \rvert$
0	15	18	0.25	0.3	0.052119	-
1	25	22	0.67	0.66	0.00404	**0.41263**
2	10	13	0.83	0.88	0.04615	0.21282
3	5	5	0.92	0.97	0.04956	0.1329
4	4	2	0.98	0.99	0.00892	0.07559
5	1	0	1	1	0.0015	0.01517

Logo, $D = 0.41263$. Comparando com a Tabela A3 da Figura 2, temos que $0.41263 > 1.36/\sqrt{60} = 0.1756$. Logo, rejeitamos a Hipótese de que os dados vêm de uma *Poisson*(1.2).

Um ótimo exemplo de como os pacotes realizam este cálculo pode ser visto em [Korosteleva, 2014], Seção 1.5.2.zz Veja também o Exemplo 6.1.1 de [Conover, 1996].

```r
### Calculando a estatistica D manualmente ###
# F dist. empirica
count_obs=c(15,25,10,5,4,1)
f_obs=count_obs/60
# F acumulada empirica
F_obs=sapply(1:6,function(i){
     return(sum(f_obs[1:i]))})
x=0:5
# F dist. teorica Poisson
f_exp=dpois(x,1.2)
# F acumulada poisson
F_exp=ppois(x,1.2)
# Maximo entre D+ e D-
D_1=unlist(sapply(1:6,function(i){
   return(try(abs(F_exp[i]-F_obs[i]),
                     silent=T))}))
D_2=unlist(sapply(1:6,function(i){
   return(try(abs(F_exp[i]-F_obs[i-1]),
                        silent=T))}))
D=max(D_1,D_2)
# [1] 0.4126273

### Calculando a estatistica D usando ks.test()
x=unlist(sapply(0:5,function(i){
        return(rep(i,count_obs[i+1]))}))
ks.test(x,"ppois",1.2)
# One-sample Kolmogorov-Smirnov test
# D = 0.41263, p-value = 2.678e-09
# alternative hypothesis: two-sided
```

Testes Unilaterais

Teste Unilateral à Esquerda:

H_0: $F(x) \geq F^*(x)$, para todo $-\infty < x < \infty$.

H_1: $F(x) < F^*(x)$ para pelo menos um valor de x.

$$D^+ = \sup_x [F^*(x) - S(x)],$$

em que D^+ corresponde a maior distância positiva entre $F^*(x)$ e $S(x)$.

Teste Unilateral à Direita:

H_0: $F(x) \leq F^*(x)$, para todo $-\infty < x < \infty$.

H_1: $F(x) > F^*(x)$ para pelo menos um valor de x.

$$D^- = \sup_x [S(x) - F^*(x)],$$

em que D^- corresponde a maior distância "negativa"entre $F^*(x)$ e $S(x)$

Considerações

- A tabela A13 da Figura 2 fornece valores exatos somente se $F^*(x)$ for contínua, com amostras menores ou iguais a 40.
- Na maioria das vezes, principalmente para pequenas amostras, o teste de Kolmogorov-Smirnov é um teste mais poderoso que o teste Qui-Quadrado.

Teste de Lilliefors

- [Lilliefors, 1967] introduziu uma modificação ao teste de Kolmogorov-smirnov para testar a normalidade de dados sem a necessidade de especificar seus parâmetros.
- Os parâmetros são estimados com os dados amostrais calculando os escores reduzidos.

Hipóteses do teste

H_0: A amostra é proveniente de uma população com distribuição normal, com média e desvio-padrão desconhecidos.

H_1: A amostra é não proveniente de uma população com distribuição normal.

Cálculo I

- Dada uma amostra aleatória $x_1, ..., x_n$ da v.a. de interesse, calcular a média amostral $\overline{x}$ e o desvio-padrão amostral S;
- Calcular os escores padronizados $z_i = \dfrac{x_i - \overline{x}}{S}, \quad i = 1, 2, \ldots, n$;
- Calcular os escores ordenados $z_{(i)}$;
- Com os escores ordenados da amostra, obtém-se a distribuição de frequência acumulada $\hat{Z}(x)$.
- Defina agora $F^*(x)$ como a probabilidade acumulada da Normal-padrão, $N(0, 1)$, até os valores $z_{(i)}$:

$$F^*(x_i) = \Phi(z_{(i)}).$$

- A estatística de teste Lilliefors T_1 é definida por:

$$T_1^+ = \sup_x |F^*(x_{(i)}) - \hat{Z}(x_{(i)})| \text{ e } T_1^- = \sup_x |F^*(x_{(i)}) - \hat{Z}(x_{(i-1)})|$$

$$T_1 = max(T_1^+, T_1^-),$$

  que corresponde a maior distância vertical entre $F^*(x)$ e $\hat{Z}(x)$.
- Compara-se o valor observado T_1 com o valor crítico tabelado para o teste de Lilliefors.
- Para um nível de significância α, rejeitar H_0 se a estatística T_1 exceder ao valor do quantil tabelado de $1 - \alpha$.

Cálculo II

TABLE A14 Quantiles of the Lilliefors Test Statistic for Normality[a]

	p = 0.80	0.85	0.90	0.95	0.99
Sample size n = 4	0.303	0.320	0.344	0.374	0.414
5	0.290	0.302	0.319	0.344	0.398
6	0.268	0.280	0.295	0.321	0.371
7	0.252	0.264	0.280	0.304	0.353
8	0.239	0.251	0.266	0.290	0.333
9	0.227	0.239	0.253	0.275	0.319
10	0.217	0.228	0.241	0.262	0.303
11	0.209	0.219	0.232	0.252	0.291
12	0.201	0.210	0.223	0.243	0.281
13	0.193	0.203	0.215	0.233	0.270
14	0.187	0.196	0.209	0.227	0.264
15	0.181	0.190	0.202	0.219	0.256
16	0.176	0.184	0.195	0.212	0.248
17	0.170	0.179	0.190	0.207	0.241
18	0.166	0.174	0.185	0.201	0.234
19	0.162	0.171	0.181	0.197	0.230
20	0.159	0.167	0.177	0.192	0.223
21	0.155	0.163	0.173	0.188	0.219
22	0.152	0.160	0.170	0.185	0.214
23	0.149	0.156	0.165	0.181	0.210
24	0.145	0.153	0.162	0.177	0.205
25	0.144	0.151	0.159	0.173	0.202
26	0.141	0.147	0.156	0.170	0.198
27	0.138	0.145	0.153	0.166	0.193
28	0.136	0.142	0.151	0.165	0.191
29	0.134	0.140	0.149	0.162	0.188
30	0.132	0.138	0.146	0.159	0.183
≥31	$\frac{0.741}{d_n}$	$\frac{0.775}{d_n}$	$\frac{0.819}{d_n}$	$\frac{0.895}{d_n}$	$\frac{1.035}{d_n}$

$d_n = (\sqrt{n} - 0.01 + 0.83/\sqrt{n})$

Figura 3: Tabela dos valores críticos para o teste de Lilliefors. Fonte: [Conover, 1996].

Exemplo

Um distribuidor pretende estimar o tempo médio de entrega dos seus produtos a um cliente importante. Foi colhida uma amostra de cinco tempos: $29, 33, 35, 36$ e 36. O senhor quer estimar o tempo médio pretendido através de um intervalo de confiança, mas não sabe nada acerca da distribuição do tempo de entrega X, e além disso, a dimensão da amostra é muito pequena ($n = 5$). Poderá fazê-lo?

```
library(nortest)
lillie.test(c(29,22,35,36,36))

# Lilliefors (Kolmogorov-Smirnov) normality test

# data:  c(29, 22, 35, 36, 36)
# D = 0.31114, p-value = 0.1147
```

Da Tabela A14 da Figura 3, temos que o valor do quantil 0.95 é 0.344. Como $D = 0.31113 < 0.344$ então não rejeitamos H_0. Logo, o distribuidor poderá utilizar um intervalo de confiança com a distribuição Normal.

Teste de Lilliefors para distribuição Exponencial

- É possível também utilizar o teste de Lilliefors para a distribuição Exponencial.
- Calcular os escores padronizados $z_i = \frac{x_i}{\overline{x}}, \quad i = 1, 2, \ldots, n$ e construir a distribuição empírica $\hat{Z}(x)$ como antes;
- A estatística de teste é definida por:

$$T_2 = \sup_x |\hat{Z}(x) - F^*(x)|,$$

com $F^*(x)$ a Fç. Dist. Acumulada da Exponencial.

Hipóteses do teste

H_0: A amostra é segue uma distribuição exponencial com $F(x) = 1 - e^{x/t}, x > 0$.

H_1: a distribuição de X não é exponencial.

- Valores tabelados para o teste bilateral em [Conover, 1996] na Tabela A15. Para o teste unilateral, veja [Durbin, 1975].

Teste de Anderson-Darling

- É uma alternativa aos testes de Qui-quadrado de aderência e Kolmogorov-Smirnov, sendo que também considera uma escala ao menos ordinal;
- Sua metodologia é próxima ao teste KS, que usa uma forma de função distância para calcular a similaridade entre $S(x)$,função distribuição empírica, e $F^*(x)$, distribuição teórica. Porém, ao contrário do KS, a estatística AD considera o intervalo total de valores dos dados, ao invés de simplesmente o máximo desvio;
- Este teste valoriza às caudas da distribuição de frequência, sendo mais recomendado para distribuições assintóticas;
- O teste de Anderson-Darling faz uso de uma distribuição específica para o cálculo de valores críticos (tabelados).

Hipóteses

H_0: A v.a. estudada segue a distribuição específica.
H_1: A amostra não é proveniente da distribuição especificada.

Estatística de Teste I

- Seja $X_1, X_2, \ldots, X_n$ a.a. com função de distribuição desconhecida, denotada por $F^*(x)$.
- $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ denotam as estatísticas de ordem da amostra aleatória.
- A estatística de teste de Anderson-Darling é definida por:

$$A = n \int_{-\infty}^{\infty} \frac{(S(x) - F^*(x))^2}{F^*(x)\,(1 - F^*(x))} dF^*(x), \tag{1}$$

  que corresponde a uma medida total de distância quadrática entre $S(x)$ e $F^*(x)$, ponderada por $[F^*(x)\,(1 - F^*(x))]^{-1}$.
- Para facilitar pode-se utilizar a simplificação proposta por [Arshad et al., 2003], calculada diretamente utilizando as estatísticas de ordem e distribuição acumulada:

$$A' = -n - \frac{1}{n} \sum_{i=1}^{n} (2i - 1) \left(\log F^*(x_{(i)}) + \log(1 - F^*(x_{(n+1-i)})) \right)$$

Estatística de Teste II

- Para amostras pequenas, a estatística computada é frequentemente multiplicada por um fator de correção, k, que varia de acordo com a distribuição que está sendo testada.
- Para o Normal, com média e desvio padrão estimados, o ajuste amplamente utilizado é:

$$k = \left(1 + \frac{4}{n} + \frac{25}{n^2}\right)$$

  sendo $A^* = k \times A'$.

- Os valores tabelados dependerão da distribuição a ser testada e do número de parâmetros a serem estimados. No caso da Normal, teremos:
  - Caso 1: média e variância conhecidos;
  - Caso 2: variância conhecida, mas média desconhecida;
  - Caso 3: média conhecida, mas variância desconhecida;
  - Caso 4: média e variância desconhecida.

Alguns valores críticos para o teste AD para a Normal são apresentados na Tabela 1.

Estatística de Teste III

Tabela 1: Tabela de Valores Críticos para a Normal

Caso	n	15%	10%	5%	2.5%	1%
1	≥ 5	1.621	1.933	2.492	3.070	3.878
2			0.908	1.105	1.304	1.573
3	≥ 5	1.760	2.323	2.904	3.690	
4	10	0.514	0.578	0.683	0.779	0.926
	20	0.528	0.591	0.704	0.815	0.969
	50	0.546	0.616	0.735	0.861	1.021
	100	0.559	0.631	0.754	0.884	1.047
	n grande	0.576	0.656	0.787	0.918	1.092

Decisão do Teste:

Para um nível de significância α, rejeitar H_0 se a estatística A ou AD exceder ao valor do quantil tabelado de $1-\alpha$.

Estatística de Teste IV

Observações:

- Quanto melhor a distribuição se ajusta aos dados, menor será a estatística A;
- O teste de Anderson-Darling é essencialmente unilateral, no sentido de que testa se existe ou não uma distância quadrática geral da distribuição observada á esperada;
- Sua principal vantagem é permitir um teste mais sensível à qualquer discrepância, e não apenas à discrepância máxima;
- A desvantagem é que os valores críticos devem ser calculados para cada distribuição;
- Existem tabelas para outras distribuições como uniforme, lognormal, exponencial, Weibull e pareto generalizada etc. Veja e.g. [Stephens, 1974] e [Koziol et al., 1987].
- Em alguns casos, o p-valor não existe matematicamente, i.e., a integral (1) não pode ser calculada.

Aspecto computacional

No R:

```
library(nortest)
ad.test()
```

No SAS:

```
proc univariate  NORMAL PLOT;
run;
proc capability;
run;
```

Exemplo I

Um especialista de qualidade pretende verificar se a distribuição dos perímetros(em cm) dos itens produzidos por sua empresa possui distribuição normal. Para isso coletou uma amostra de 5 itens: $12.29, 11.40, 14.22, 12.37, 11.91$. Verifique esta hipótese utilizando o teste AD considerando 5% de significância.

i	$x_{(i)}$	$F(x_{(i)})$	$F(x_{(n+1-i)})$	$logF(x_{(i)})$	$log(1-F(x_{(n+1-i)}))$
1	11.40	0.15	0.96	-1.91	-3.29
2	11.91	0.30	0.47	-1.20	-0.64
3	12.29	0.44	0.44	-0.82	-0.58
4	12.37	0.47	0.30	-0.75	-0.36
5	14.22	0.96	0.15	-0.04	-0.16

$$1 \cdot \left(\log F(x_{(1)}) + \log(1 - F(x_{(5)}))\right) + \ldots + 9 \cdot \left(\log F(x_{(5)}) + \log(1 - F(x_{(1)}))\right) = -27.25$$

Exemplo II

$$
\begin{aligned}
A' &= -5 - \frac{-27.25}{5} = 0.451 \\
A^* &= \left(1 + \frac{4}{5} + \frac{25}{5^2}\right) \times 0.451 = 1.26
\end{aligned}
$$

- Como tanto a média quanto a variância são desconhecidas, utilizaremos o valor critico do caso 4, e a linha com n mais próximo de 5 na Tabela 1 é $n = 10$.
- Portanto, ao nível de significância 0.05 temos que $A^* > 0.683$ e rejeitamos a hipótese dos dados serem provenientes de uma distribuição normal.

Exemplo no SAS

Concentrações de ozônio (O_3, $\mu g/m^3$) na estação meteorológica de Pinheiros na cidade de São Paulo entre 24-Set-2020 e 28-Set-2020 (dados horários, fonte: CETESB).

```
data O3_pinheiros;
label O3 = "Concentracao Horaria de Ozonio (ug/m3)";
input O3 @@;
datalines;
15 23 27 19 21 16 12 13 17 26 37 57 81 99 106 75 70 71 56 31 31 23 6 1 2 1
0 0 0 0 0 1 11 28 54 86 122 142 144 140 89 50 46 22 1 1 0 0 0 0 4 21 49 88
104 102 101 103 102 95 38 15 1 0 0 0 1 10 47 78 95 102 97 106 109 102 95 81
39 14 2 1 1 26 34 27 50 69 58 18 1 6 11 33 73  61 65 57 56 28 26
;
proc univariate NORMAL PLOT;
var O3;
run;
```

Tests for Normality

Test	Statistic		p Value	
Shapiro-Wilk	W	0.889003	Pr < W	<0.0001
Kolmogorov-Smirnov	D	0.151885	Pr > D	<0.0100
Cramer-von Mises	W-Sq	0.59487	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	3.710862	Pr > A-Sq	<0.0050

Teste de Shapiro-Wilk

- Em contraste com outros testes de aderência por comparação de distribuições teórica e observada, o teste de Shapiro-Wilk só é aplicável para verificar a normalidade;
- O teste a variância amostral com a variância estimada a partir da distribuição amostral das estatísticas de ordem;
- Possui robustez para detectar fuga de normalidade por assimetria ou curtose (ou ambos);
- Não exige especificação dos parâmetros da distribuição normal a ser testada, apenas qua a amostra seja aleatória.

Hipóteses

H_0: A variável aleatória amostrada possui distribuição normal com média e variância não especificadas.

H_1: A v.a. amostrada não é normal.

Estatística de Teste I

- Seja $X_1, X_2, \ldots, X_n$ amostra aleatória com função de distribuição desconhecida, denotada por $F(x)$, mensurada na escala ao menos ordinal;
- $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ denotam as estatísticas de ordem da amostra aleatória;
- A estatística de teste de Shapiro-Wilk é definida por:

$$W = \frac{\left[\sum_{i=1}^{k} a_i \left(x_{(n+1-i)} - x_{(i)}\right)\right]^2}{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}$$

em que

$$k = \begin{cases} (n+1)/2 & \text{se } n \text{ impar,} \\ n/2 & \text{se } n \text{ par,} \end{cases}$$

e a_i obtidos pela Tabela A16 [Conover, 1996].

- O vetor $\mathbf{a} = (a_1, ..., a_k)'$ é função da esperança e variância das estatísticas de ordem de amostras de tamanho n da distribuição normal padrão;

Estatística de Teste II

- A estatística W está entre 0 e 1; W pequeno é evidência de fuga de normalidade; valores próximos a 1 indicam normalidade.

Decisão do Teste:

Para um nível de significância α, rejeitar H_0 se a estatística W for menor que o quantil tabelado de α. Tabela A17 de [Conover, 1996].

Observações

- Proposto inicialmente por [Shapiro and Wilk, 1965] para $n \leq 50$ e posteriormente foi aumentado para n até 5000 [Royston, 1982];
- Os quantis (tabelados) para a estatística W foram obtidos por meio de simulações de Monte Carlo por [Pearson and Hartley, 1972];
- Este teste é essencialmente unilateral a esquerda;
- Eficiente mesmo com um número pequeno de observações;
- Baixo poder na existência de muitos empates;

Aspecto Computacional

No R:

```
library(nortest)
shapiro.test()
```

No SAS:

```
proc univariate NORMAL PLOT;
run;
```

Acknowledgements

Agradecemos ao prof. Anderson Ara pela disponibilização de seu material didático, no qual nos baseamos para a elaboração destes slides. Alguns trechos desta apresentação são replicados de seu material.

Referências I

Arshad, M., Rasool, M. T., and Ahmad, M. I. (2003).
Anderson darling and modified anderson darling tests for generalized pareto distribution.
Journal of Applied Sciences, 3:85–88.

Conover, W. J. (1996).
Practical nonparametric statistics.
John Wiley and sons, 3 ed. edition.

Durbin, J. (1975).
Kolmogorov-smirnov tests when parameters are estimated with applications to tests of exponentiality and tests on spacings.
Biometrika, 62:5–22.

Korosteleva, O. (2014).
Nonparametric methods in statistics with SAS applications.
Crc Press.

Referências II

Koziol, J. A., D'Agostino, R. B., and Stephens, M. A. (1987).
Goodness-of-fit techniques.
Journal of Educational Statistics, 12:412.

Lilliefors, H. W. (1967).
On the kolmogorov-smirnov test for normality with mean and variance unknown.
Journal of the American Statistical Association, 62:399–402.

Pearson, A. V. and Hartley, H. O. (1972).
Biometrica Tables for Statisticians, volume 2.
Cambridge University Press.

Royston, J. P. (1982).
An extension of shapiro and wilk's w test for normality to large samples.
Journal of the Royal Statistical Society. Series C (Applied Statistics), 31:115–224.

Referências III

Shapiro, S. S. and Wilk, B. M. (1965).
An analysis of variance test for normality (complete samples).
Biometrika, 52:591–611.

Stephens, M. A. (1974).
Edf statistics for goodness of fit and some comparisons.
Journal of the American Statistical Association, 69:730–737.