

Modelling the dependence between two diagnostic tests via copula functions

Jorge Alberto Achcar¹

José Rafael Tovar Cuevas²

¹Department of Social Medicine FMRP, University of São Paulo,
Ribeirão Preto, SP, Brazil

²Department of Statistics, IMECC, State University of Campinas,
Campinas, SP, Brazil

Abstract

In medical diagnostic testing, it is common the use of more than one diagnostic test applied to the same individual. Usually these tests are assumed to be independents and important performance measures are estimated as the sensitivities and specificities of the tests, in the presence or not of a reference test usually known as "gold standard". These tests could be dependent since they are applied to the same individual and this assumption could modify the estimation of the performance measures. Considering two diagnostic tests, we could assume a bivariate Bernoulli distribution. Alternatively, we propose the use of different copula functions to model the association between tests. Under the Bayesian paradigm, the posterior summaries of interest are obtained using MCMC (Markov Chain Monte Carlo) methods. A detailed discussion on the elicitation of prior distributions on the test performance and copula parameter is considered in this study. We illustrate the proposed methodology considering two medical data sets introduced in the literature.

Desempenho de testes diagnósticos clínicos

O diagnóstico clínico para indivíduos suspeitos de ter alguma doença é feito avaliando-se no organismo dos mesmos, algumas características biológicas que se vêm alteradas ante a presença do evento externo (doença ou infecção). O objetivo principal é estimar os parâmetros que identificam as características de eficiência dos testes (pode ser um ou mais) usados para a classificação dos indivíduos como:

- (a) **Sensibilidade,**
- (b) **Especificidade,**
- (c) **Prevalência.**

Na população de interesse, é possível encontrar duas sub-populações, uma composta por indivíduos doentes e outra composta por indivíduos não doentes. Seja D uma variável aleatória que identifica o verdadeiro estado de saúde de um indivíduo, de modo que $D = 1$ denota o caso de um doente e $D = 0$ denota um caso de não doente.

Seja também $p = P(D = 1)$ a probabilidade de encontrar um indivíduo doente na população a qual se define como **prevalência**.

Quando se tem estudos de diagnóstico clínico com dois testes de triagem e um padrão-ouro, é possível definir duas variáveis aleatórias T_1 e T_2 que identificam os resultados dos testes num indivíduo, de modo que $T_v = 1$, identifica um resultado positivo e $T_v = 0$ identifica um resultado negativo no teste $v = 1, 2$. Definir a **sensibilidade e especificidade** do teste v como:

$$S_v = P(T_v = 1 \mid D = 1) \text{ (sensibilidade)}$$

$$E_v = P(T_v = 0 \mid D = 0) \text{ (especificidade)}$$

para $v = 1, 2$.

Outros índices de desempenho para testes diagnósticos de importância em saúde pública e na tomada de decisões são: o **valor preditivo positivo (VPP)** que é definido como a probabilidade preditiva de encontrar um caso positivo na população dado por $P(D = 1 \mid T_v = 1)$ e o **valor preditivo negativo**

(VPN) definido como a probabilidade preditiva de encontrar um não doente na população dado por $P(D = 0 | T_v = 0)$ para $v = 1, 2$.

O **viés de verificação** pode se apresentar de diferentes formas dependendo do planejamento estabelecido para o estudo de avaliação de testes. Tem-se viés de verificação quando é preciso avaliar o desempenho de um teste novo e não se tem um padrão-ouro para comparação ou quando só uma parte dos indivíduos avaliados com o teste novo é submetida a verificação por padrão-ouro. Uma das formas mais comuns de viés de verificação é a que acontece quando se tem estudos com dois testes de triagem e só se verificam com padrão-ouro os indivíduos com resultado positivo em um dos dois testes em estudo, de modo que aqueles que tem ambos os resultados negativos não são verificados.

Algumas referencias: Staquet et al. (1981), Begg (1988), Blackstone e Lauer (2004), Whiting (2004) e Gupta e Roehrborn (2004) entre outros.

Quando consideramos dois testes diagnósticos aplicados ao mesmo indivíduo, poderíamos ter uma estrutura de dependência para os testes o que pode afetar a estimação das medidas de desempenho dos testes. Este problema assumindo uma estrutura binária para os testes foi estudada por diferentes autores (ver por exemplo, Thibodeau, 1981; Vacek, 1985). Quando não temos todos indivíduos verificados por um padrão ouro ou "gold standard", por exemplo, quando só os indivíduos com pelo menos um resultado positivo são verificados por este teste de referencia, poderíamos assumir modelos de variáveis latentes ou de efeitos aleatórios (ver por exemplo, Baker, 1995; Qu et al. 1996, 1998; Torrance-Rynard e Walter, 1997; Yang e Becker, 1997; Hui e Zhou, 1998, Zhou, 1998; Albert et al. 2001; Pepe e Alonso, 2001; Garret et al. 2002). Outros estudos usando métodos Bayesianos são introduzidos na literatura (Dendukuri et al, 2001; Achcar et al, 2005; Martinez et al, 2005, 2006, 2008, 2009).

Procedimentos de diagnóstico clínico usando dois ou mais testes de triagem e um padrão ouro

Em muitas situações o procedimento diagnóstico exige a aplicação de dois testes diagnósticos sob avaliação e a verificação com um padrão-ouro.

É possível que a verificação seja feita ou não para todos os participantes do estudo. Em casos onde só se verifica uma parte dos indivíduos envolvidos na avaliação, diz-se que o estudo tem viés de verificação. Aqui considera-se a aplicação de dois ou três testes de triagem com posterior verificação usando um padrão-ouro só em indivíduos que tem resposta positiva em pelo menos um dos testes de triagem.

Tabela 2.2: Arranjo para os resultados de dois testes de diagnóstico clínico e um padrão-ouro. Os valores entre colchetes são desconhecidos quando o planejamento tem viés de verificação

	Indivíduos doentes			Indivíduos não doentes		
	T_2+	T_2-	Total	T_2+	T_2-	Total
T_1+	r_1	r_2	$r_1 + r_2$	r_5	r_6	$r_5 + r_6$
T_1-	r_3	$[r_4]$	$r_3 + [r_4]$	r_7	$[r_8]$	$r_7 + [r_8]$
Total	$r_1 + r_3$	$r_2 + [r_4]$	$[n_+]$	$r_5 + r_7$	$r_6 + [r_8]$	$[n_-]$

Modelo 1: testes condicionalmente independentes

Prevalencia: $p = P(D = 1)$.

Sensibilidade: $S_v = P(T_v = 1 \mid D = 1)$

Especificidade: $E_v = P(T_v = 0 \mid D = 0)$ para teste $v = 1, 2$.

Assim,

$$P(T_1 = 1, T_2 = 1, D = 1) = P(D = 1)P(T_1 = 1 \mid D = 1)P(T_2 = 1 \mid T_1 = 1, D = 1),$$

then, assuming independence between tests we have,

$$P(T_1 = 1, T_2 = 1, D = 1) = P(D = 1)P(T_1 = 1 \mid D = 1)P(T_2 = 1 \mid D = 1) = pS_1S_2.$$

Similarmente, obtém-se as probabilidades conjuntas para os outros casos .

Table 1: Likelihood contributions of all possible combinations of outcomes of T_1 , T_2 and D . (Values in brackets may be unknown, f_i = number of individuals in the cell $i = 1, 2, \dots, 8$)

					Contribution to likelihood	
i	D	T_1	T_2	f_i	Two independent tests	
1	1	1	1	f_1	pS_1S_2	
2	1	1	0	f_2	$pS_1(1 - S_2)$	
3	1	0	1	f_3	$p(1 - S_1)S_2$	
4	1	0	0	$[f_4]$	$p(1 - S_1)(1 - S_2)$	
5	0	1	1	f_5	$(1 - p)(1 - E_1)(1 - E_2)$	
6	0	1	0	f_6	$(1 - p)(1 - E_1)E_2$	
7	0	0	1	f_7	$(1 - p)E_1(1 - E_2)$	
8	0	0	0	$[f_8]$	$(1 - p)E_1E_2$	

Definir duas variáveis aleatórias Y_1 e Y_2 como os números de indivíduos doentes e não-doentes entre as pessoas não verificadas. As variáveis aleatórias Y_1 e Y_2 respectivamente, são definidas como variáveis latentes (ver Tanner e Wong, 1987) e $Y_2 = u - Y_1$, onde u é o número de indivíduos não verificados. Observar que f_4 and f_8 são valores desconhecidos, mas a soma $u = f_4 + f_8$ conhecida. Podemos simular Y_1 de uma distribuição binomial dada por,

$$Y_1|u, \theta_1 \sim b \left\{ u; \frac{p(1 - S_1)(1 - S_2)}{p(1 - S_1)(1 - S_2) + (1 - p)E_1E_2} \right\},$$

onde $b(u; \theta)$ denota uma distribuição binomial com média $u\theta$ e variância $u\theta(1-\theta)$;

$$\theta = \frac{p(1 - S_1)(1 - S_2)}{p(1 - S_1)(1 - S_2) + (1 - p)E_1E_2}$$

$$\theta_1 = [\theta_{1,1} = S_1, \theta_{1,2} = S_2, \theta_{1,3} = E_1, \theta_{1,4} = E_2, \theta_{1,5} = p]$$

Assumir distribuições Beta (α, β) (independencia a priori) a priori para os cinco parâmetros.

Distribuições condicionais para o amostrador de Gibbs:
distribuições Beta para todos os parâmetros:

Vetor de quantidades observadas dado por $Z=[a,b,c,d,e,f,g]$:

$$p|Z, y_1, y_2, \alpha_p, \beta_p \sim \text{Beta}(a + b + c + y_1 + \alpha_p, e + f + g + y_2 + \beta_p),$$

$$S_1|Z, y_1, y_2, \alpha_{S_1}, \beta_{S_1} \sim \text{Beta}(a + b + \alpha_{S_1}, c + y_1 + \beta_{S_1}),$$

$$S_2|Z, y_1, y_2, \alpha_{S_2}, \beta_{S_2} \sim \text{Beta}(a + c + \alpha_{S_2}, b + y_1 + \beta_{S_2}),$$

$$E_1|Z, y_1, y_2, \alpha_{E_1}, \beta_{E_1} \sim \text{Beta}(g + y_2 + \alpha_{E_1}, e + f + \beta_{E_1}),$$

$$E_2|Z, y_1, y_2, \alpha_{E_2}, \beta_{E_2} \sim \text{Beta}(f + y_2 + \alpha_{E_2}, e + g + \beta_{E_2}).$$

Uso de funções cópulas

Quando se tem planejamentos de estudos que incluem a medição de duas ou mais variáveis aleatórias, uma hipótese de interesse é a existência de dependência entre as variáveis, o que faz com que seja muito importante procurar por uma medida de dependência apropriada para os dados.

Uma cópula é uma função de distribuição conjunta de variáveis aleatórias com distribuição uniforme padrão $U(0,1)$, isto é:

$$C(u_1, \dots, u_d) = P(U_1 \leq u_1, \dots, U_d \leq u_d)$$

onde $U_i \sim U(0, 1)$ para $i = 1, \dots, d$; então as funções de cópula permitem caracterizar a estrutura de dependência de um conjunto

de variáveis aleatórias independentemente das distribuições marginais.

Dadas X_1, \dots, X_d variáveis aleatórias com distribuição de probabilidade conjunta F e funções de distribuição marginais $F_i, i = 1, \dots, d$, Sklar (1959) mostrou que, para alguma distribuição multivariada F , existe uma única função de cópula C que pode ser escrita como,

$$C(u_1, \dots, u_d) = F[F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)]$$

De outro lado, é fácil mostrar que, se C é uma função cópula e F_1, \dots, F_d são funções de distribuição arbitrárias, então a função F definida como

$$F(x_1, \dots, x_d) = C[F_1(x_1), \dots, F_d(x_d)]$$

é uma função de distribuição multivariada com funções de distribuição marginais F_1, \dots, F_d .

Assumir que os testes diagnósticos são realizações das variáveis aleatórias V_1 e V_2 medidas numa escala contínua positiva, isto é, $V_1 > 0$ e $V_2 > 0$. Assumir os pontos de corte ξ_1 e ξ_2 para cada teste tal que o resultado é positivo se $V_v > \xi_v$, isto é, $T_v = 1$ se e só se $V_v > \xi_v$ para $v=1,2$.

Modelo 2: cópula de Farlie-Gumbel-Morgenstern (FGM) definida por,

$$C_I(u, w) = uw[1 + \varphi(1 - u)(1 - w)],$$

onde $u = F_1(v_1)$, $w = F_2(v_2)$ e $-1 \leq \varphi \leq 1$ (apropriada para dependências fracas) e φ mede a dependência entre as duas marginais, de modo que, se $\varphi = 0$, as duas variáveis aleatórias são independentes. Assume-se dois parâmetros de dependência distintos: φ_D e φ_{ND} para doentes e não doentes respectivamente.

O parâmetro φ é relacionado com os coeficientes de concordância Tau de Kendall (ζ) e Rho de Spearman (ρ) pelas equações:

$$\tau_{v,w} = 4 \int \int C(u, w) dC(u, w) - 1 = 4 \left(\frac{\varphi}{18} + \frac{1}{4} \right) - 1 = \frac{2\varphi}{9},$$

$$\rho_{v,w} = 12 \int \int C(u, w) dC(u, w) - 3 = 12 \left(\frac{1}{4} + \frac{\varphi}{36} \right) - 3 = \frac{\varphi}{3}.$$

Assim a função de distribuição acumulada e a função de sobrevivência são dadas por,

$$F(\xi_1, \xi_2) = C(F_1(\xi_1), F_2(\xi_2)) = F_1(\xi_1)F_2(\xi_2)[1 + \varphi(1 - F_1(\xi_1))(1 - F_2(\xi_2))],$$

$$S(\xi_1, \xi_2) = P(V_1 > \xi_1, V_2 > \xi_2) = 1 - F_1(\xi_1) - F_2(\xi_2) + F(\xi_1, \xi_2).$$

Contribuições à verossimilhança em termos da função de cópula quando tem-se dois testes de triagem e um padrão-ouro:

$$P(T_1 = 1|D = 1) = P(V_1 > \xi_1|D = 1) = S_1,$$

$$\text{e, } P(T_2 = 1|D = 1) = P(V_2 > \xi_2|D = 1) = S_2,$$

$$\text{então } P(T_1 = 0|D = 1) = P(V_1 \leq \xi_1|D = 1) = F_1^D(\xi_1) = 1 - S_1$$

$$\text{e, } P(T_2 = 0|D = 1) = P(V_2 \leq \xi_2|D = 1) = F_2^D(\xi_2) = 1 - S_2.$$

$$\begin{aligned} \text{de modo que, } F_D(\xi_1, \xi_2) &= P(V_1 \leq \xi_1, V_2 \leq \xi_2|D = 1), \\ &= F_1^D(\xi_1)F_2^D(\xi_2)[1 + \varphi(1 - F_1^D(\xi_1))(1 - F_2^D(\xi_2))], \\ &= (1 - S_1)(1 - S_2)(1 + \varphi_D S_1 S_2). \end{aligned}$$

$$\begin{aligned} \text{Se, } P(T_1 = 1, T_2 = 1|D = 1) &= P(V_1 > \xi_1, V_2 > \xi_2|D = 1), \\ &= S_D(\xi_1, \xi_2), \\ &= 1 - F_1^D(\xi_1) - F_2^D(\xi_2) + F_D(\xi_1, \xi_2), \\ &= 1 - (1 - S_1) - (1 - S_2) + (1 - S_1)(1 - S_2)(1 + \varphi_D S_1 S_2), \\ &= S_1 S_2 [1 + \varphi_D (1 - S_1)(1 - S_2)]. \end{aligned}$$

$$\begin{aligned} \text{Então, } P(T_1 = 1, T_2 = 1, D = 1) &= P(D = 1)P(T_1 = 1, T_2 = 1|D = 1), \\ &= p S_1 S_2 [1 + \varphi_D (1 - S_1)(1 - S_2)]. \end{aligned}$$

Para os outros subgrupos de indivíduos doentes, se tem que:

$$\begin{aligned}
P(T_1 = 1, T_2 = 0, D = 1) &= P(D = 1)P(T_1 = 1, T_2 = 0|D = 1), \\
&= pP(V_1 > \xi_1, V_2 \leq \xi_2|D = 1), \\
&= pP(V_2 \leq \xi_2|D = 1) - P(V_1 \leq \xi_1, V_2 \leq \xi_2|D = 1), \\
&= p[F_2^D(\xi_2) - F_D(\xi_1, \xi_2)], \\
&= p[F_2^D(\xi_2) - F_1^D(\xi_1)F_2^D(\xi_2)[1 + \varphi_D(1 - F_1(\xi_1)^D)(1 - F_2^D(\xi_2))]], \\
&= p[(1 - S_2) - (1 - S_1)(1 - S_2)(1 + \varphi_S S_1 S_2)], \\
&= p(1 - S_2)[S_1 - \varphi_D S_1 S_2(1 - S_1)].
\end{aligned}$$

$$\begin{aligned}
P(T_1 = 0, T_2 = 1, D = 1) &= P(D = 1)P(T_1 = 0, T_2 = 1|D = 1), \\
&= pP(V_1 \leq \xi_1, V_2 > \xi_2|D = 1), \\
&= P(V_1 \leq \xi_1|D = 1) - P(V_1 \leq \xi_1, V_2 \leq \xi_2|D = 1), \\
&= F_1^D(\xi_1) - F_D(\xi_1, \xi_2), \\
&= p[F_1^D(\xi_1) - F_1^D(\xi_1)F_2^D(\xi_2)[1 + \varphi_D(1 - F_1(\xi_1)^D)(1 - F_2^D(\xi_2))]], \\
&= p[(1 - S_1) - (1 - S_1)(1 - S_2)(1 + \varphi_S S_1 S_2)], \\
&= p(1 - S_1)[S_2 - \varphi_D S_1 S_2(1 - S_2)],
\end{aligned}$$

$$\begin{aligned}
P(T_1 = 0, T_2 = 0, D = 1) &= P(D = 1)P(T_1 = 0, T_2 = 0|D = 1), \\
&= pP(V_1 \leq \xi_1, V_2 \leq \xi_2|D = 1), \\
&= pF_D(\xi_1, \xi_2), \\
&= p(1 - S_1)(1 - S_2)[1 + \varphi_D S_1 S_2].
\end{aligned}$$

No grupo de indivíduos com padrão-ouro negativo os resultados são:

$$P(T_1 = 0|D = 0) = P(V_1 \leq \xi_1|D = 0) = F_1^{ND}(\xi_1) = E_1,$$

e, $P(T_2 = 0|D = 0) = P(V_2 \leq \xi_2|D = 0) = F_2^{ND}(\xi_2) = E_2.$

Então, $P(T_1 = 1, T_2 = 1|D = 0) = P(V_1 > \xi_1, V_2 > \xi_2|D = 0),$

$$= S_{ND}(\xi_1, \xi_2),$$

$$= (1 - F_1^{ND}(\xi_1) - F_2^{ND}(\xi_2) + F_{ND}(\xi_1, \xi_2)),$$

$$= (1 - E_1)(1 - E_2)[1 + \varphi_{ND}E_1E_2].$$

e, $P(T_1 = 1, T_2 = 1, D = 0) = (1 - p)(1 - E_1)(1 - E_2)[1 + \varphi_{ND}E_1E_2].$

Finalmente, $P(T_1 = 0, T_2 = 0|D = 0) = P(V_1 \leq \xi_1, V_2 \leq \xi_2|D = 0),$

$$= F_{ND}(\xi_1, \xi_2),$$

$$= E_1E_2[1 + \varphi_{ND}(1 - E_1)(1 - E_2)].$$

Portanto, $P(T_1 = 0, T_2 = 0, D = 0) = (1 - p)E_1E_2[1 + \varphi_{ND}(1 - E_1)(1 - E_2)].$

Continuando com o esse procedimento, obtemos todas as contribuições à função de verossimilhança em estudos de diagnóstico clínico com dois testes de triagem e um padrão-ouro, as quais aparecem na tabela abaixo:

Tabela 4.5: Contribuições à função de verossimilhança das combinações de resultados entre T_1 , T_2 e D em planejamentos que assumem dependência entre dois testes de resposta com estrutura de dependência FGM. Os valores entre colchetes são não observáveis em presença de viés de verificação. (r_i , $i = 1, 2, \dots, 8$ é o número de indivíduos em cada célula da tabela)

D	T_1	T_2	r_i	Contribuição à verossimilhança
1	1	1	r_1	$pS_1S_2[1 + \varphi_D(1 - S_1)(1 - S_2)]$
1	1	0	r_2	$pS_1(1 - S_2)[1 - \varphi_D(1 - S_1)S_2]$
1	0	1	r_3	$p(1 - S_1)S_2[1 - \varphi_DS_1(1 - S_2)]$
1	0	0	$[r_4]$	$p(1 - S_1)(1 - S_2)[1 + \varphi_DS_1S_2]$
0	1	1	r_5	$(1 - p)(1 - E_1)(1 - E_2)[1 + \varphi_{ND}E_1E_2]$
0	1	0	r_6	$(1 - p)(1 - E_1)E_2[1 - \varphi_{ND}E_1(1 - E_2)]$
0	0	1	r_7	$(1 - p)E_1(1 - E_2)[1 - \varphi_{ND}E_2(1 - E_1)]$
0	0	0	$[r_8]$	$(1 - p)E_1E_2[1 + \varphi_{ND}(1 - E_1)(1 - E_2)]$

Pode-se observar que, para todas as combinações de resultados, quando $\varphi = 0$ obtêm-se os resultados apresentados na Tabela 1, (o caso de independência entre as respostas dos testes).

Modelo 3: cópula de Gumbel definida por,

$$C(u, w) = u + w - 1 + (1 - u)(1 - w) \exp\{-\phi \ln(1 - u) \ln(1 - w)\}..$$

Para este modelo, as funções de probabilidade conjunta acumulada e de sobrevivência tem as formas:

$$\begin{aligned}
F(\xi_1, \xi_2) &= F_1(\xi_1) + F_2(\xi_2) - 1 + (1 - F_1(\xi_1))(1 - F_2(\xi_2)), \\
&\quad \exp\{-\phi \ln(1 - F_1(\xi_1)) \ln(1 - F_2(\xi_2))\}. \quad e, \\
S(\xi_1, \xi_2) &= (1 - F_1(\xi_1))(1 - F_2(\xi_2)) \exp\{-\phi \ln(1 - F_1(\xi_1)) \ln(1 - F_2(\xi_2))\}.
\end{aligned}$$

Tabela 4.7: Contribuições à função de verossimilhança das combinações de resultados entre T_1 , T_2 e D em planejamentos que assumem dependência entre dois testes de resposta com estrutura de dependência Gumbel. Os valores entre colchetes são não observáveis em presença de viés de verificação. (r_i , $i = 1, 2, \dots, 8$ é o número de indivíduos em cada célula da tabela)

D	T_1	T_2	r_i	Contribuição à verossimilhança
1	1	1	r_1	$pS_1S_2 \exp(-\phi_D \log S_1 \log S_2)$
1	1	0	r_2	$pS_1[1 - S_2 \exp(-\phi_D \log S_1 \log S_2)]$
1	0	1	r_3	$pS_2[1 - S_1 \exp(-\phi_D \log S_1 \log S_2)]$
1	0	0	$[r_4]$	$p[1 - S_1 - S_2 + S_1S_2 \exp(-\phi_D \log S_1 \log S_2)]$
0	1	1	r_5	$(1-p)(1-E_1)(1-E_2) \exp(-\phi_{ND} \log(1-E_1) \log(1-E_2))$
0	1	0	r_6	$(1-p)(1-E_1)[1 - (1-E_2) \exp(-\phi_{ND} \log(1-E_1) \log(1-E_2))]$
0	0	1	r_7	$(1-p)(1-E_2)[1 - (1-E_1) \exp(-\phi_{ND} \log(1-E_1) \log(1-E_2))]$
0	0	0	$[r_8]$	$(1-p)[E_1 + E_2 - 1 + (1-E_1)(1-E_2) \exp(-\phi_{ND} \log(1-E_1) \log(1-E_2))]$

Modelo 4: cópula de Clayton definida por,

$$\begin{aligned}
C(u, w) &= \max \left[\frac{uw}{[u^\alpha + w^\alpha - u^\alpha w^\alpha]^{1/\alpha}}, 0 \right], \quad u, w \in (0, 1), \quad \alpha \in [-1, \infty) \setminus \{0\} \\
F(\xi_1, \xi_2) &= \frac{F_1(\xi_1)F_2(\xi_2)}{[(F_1(\xi_1))^\alpha + (F_2(\xi_2))^\alpha - (F_1(\xi_1))^\alpha (F_2(\xi_2))^\alpha]^{1/\alpha}}. \\
S(\xi_1, \xi_2) &= (1 - F_1(\xi_1))(1 - F_2(\xi_2)) + F(\xi_1, \xi_2).
\end{aligned}$$

Tabela 4.9: Contribuições à função de verossimilhança das combinações de resultados entre T_1 , T_2 e D em planejamentos que assumem dependência entre dois testes de resposta com estrutura de dependência Clayton. Os valores entre colchetes são não observáveis em presença de viés de verificação. (r_i , $i = 1, 2, \dots, 8$ é o número de indivíduos em cada célula da tabela)

D	T_1	T_2	r_i	Contribuição à verossimilhança
1	1	1	r_1	$p [S_1 + S_2 - 1 + [(1 - S_1)^\alpha + (1 - S_2)^\alpha - (1 - S_1)^\alpha(1 - S_2)^\alpha]^{-1/\alpha}(1 - S_1)(1 - S_2)]$
1	1	0	r_2	$p [(1 - S_2)(1 - [(1 - S_1)^\alpha + (1 - S_2)^\alpha - (1 - S_1)^\alpha(1 - S_2)^\alpha]^{-1/\alpha}(1 - S_1))]$
1	0	1	r_3	$p [(1 - S_1)(1 - [(1 - S_1)^\alpha + (1 - S_2)^\alpha - (1 - S_1)^\alpha(1 - S_2)^\alpha]^{-1/\alpha}(1 - S_2))]$
1	0	0	$[r_4]$	$p [[(1 - S_1)^\alpha + (1 - S_2)^\alpha - (1 - S_1)^\alpha(1 - S_2)^\alpha]^{-1/\alpha}(1 - S_1)(1 - S_2)]$
0	1	1	r_5	$(1 - p) [1 - E_1 - E_2 + [E_1^\alpha + E_2^\alpha - E_1^\alpha E_2^\alpha]^{-1/\alpha} E_1 E_2]$
0	1	0	r_6	$(1 - p) [E_2(1 - [E_1^\alpha + E_2^\alpha - E_1^\alpha E_2^\alpha]^{-1/\alpha} E_1)]$
0	0	1	r_7	$(1 - p) [E_1(1 - [E_1^\alpha + E_2^\alpha - E_1^\alpha E_2^\alpha]^{-1/\alpha} E_2)]$
0	0	0	$[r_8]$	$(1 - p) [[E_1^\alpha + E_2^\alpha - E_1^\alpha E_2^\alpha]^{-1/\alpha} E_1 E_2]$

Outra Possibilidade: distribuição Bernoulli bivariada com covariâncias ψ_D e ψ_{ND}

Tabela 4.3: Contribuições à função de verossimilhança das combinações de resultados entre T_1 , T_2 e D em planejamentos que assumem dependência entre dois testes de resposta binária. Os valores entre colchetes são não observáveis em presença de viés de verificação. (r_i , $i = 1, 2, \dots, 8$ é o número de indivíduos em cada célula da tabela)

D	T_1	T_2	r_i	contribuições à verossimilhança
1	1	1	r_1	$p[S_1 S_2 + \psi_D]$
1	1	0	r_2	$p[S_1(1 - S_2) - \psi_D]$
1	0	1	r_3	$p[(1 - S_1)S_2 - \psi_D]$
1	0	0	$[r_4]$	$p[(1 - S_1)(1 - S_2) + \psi_D]$
0	1	1	r_5	$(1 - p)[(1 - E_1)(1 - E_2) + \psi_{ND}]$
0	1	0	r_6	$(1 - p)[(1 - E_1)E_2 - \psi_{ND}]$
0	0	1	r_7	$(1 - p)[E_1(1 - E_2) - \psi_{ND}]$
0	0	0	$[r_8]$	$(1 - p)[E_1 E_2 + \psi_{ND}]$

Em casos de dependência binária, tem-se sete parâmetros de interesse (duas sensibilidades, duas especificidades, uma prevalência e duas covariâncias) mas a tabela de probabilidades observadas só tem seis componentes de informação o qual implica um problema de falta de identificabilidade no modelo de estimação. Sob a perspectiva frequentista, autores como Vacek (1985) contornaram o problema fixando um subconjunto dos parâmetros a estimar assumindo-os já conhecidos.

Problema: como obter alguma informação sobre o parâmetro de dependência para desenvolver uma análise Bayesiana?

Algumas possibilidades:

- (1) Usar a relação entre os coeficientes de dependência com os coeficientes de concordância Tau de Kendall (ζ) e Rho de Spearman (ρ) .
- (2) Usar alguns índices introduzidos na literatura para medir dependências entre testes diagnósticos (ver por exemplo, Böhning e Patilea (2008):

$$\lambda_i = \frac{P(T_1 = 1|T_2 = 1, D = i)}{P(T_1 = 1, D = i)} = \frac{P(T_1 = 1, T_2 = 1, D = i)}{P(T_1 = 1, D = i)P(T_2 = 1, D = i)} \quad \lambda_i \in (0, \infty). \quad (5.1)$$

Se $\lambda_i = 1$, os testes apresentam resultados independentes; se $\lambda_i < 1$, a dependência entre testes é negativa e se $\lambda_i > 1$, então a dependência é positiva.

Por outro lado;

$$\delta_i = \frac{P(T_1 = 1, T_2 = 1, D = i)P(T_1 = 0, T_2 = 0, D = i)}{P(T_1 = 1, T_2 = 0, D = i)P(T_1 = 0, T_2 = 1, D = i)} \quad \delta_i \in (0, \infty). \quad (5.2)$$

De modo que δ_i pode ser visto como o odds ratio no i -ésimo estado de saúde (neste caso dois) e quando $\delta_i = 1$ tem-se independência entre testes; dependências negativas são expressas por $\delta_i < 1$ e dependências positivas por $\delta_i > 1$.

Daí é possível achar relações entre os parâmetros de desempenho e prevalência com esses índices.

Algumas considerações para elicitar as distribuições a priori

A elicitação da distribuição a priori é um dos mais importantes passos numa análise de dados sob o enfoque Bayesiano. Nesta etapa do estudo, o estatístico busca um especialista ou fontes externas (artigos, relatórios, livros, bancos de dados) para obter informação sobre o parâmetro (ou parâmetros) de interesse e depois expressá-la em forma de um modelo de probabilidade que será acrescentado à função de verossimilhança para obter as distribuições a posteriori de interesse.

The problem of identifiability has been extensively discussed by many authors under the Bayesian paradigm (see for example, Lindley, 1971; Dawid, 1979; Paulino and Pereira, 1994). Poirier (1998) points out that “A

Bayesian analysis of a nonidentified model is always possible if a proper prior on all the parameters is specified” (see also, Eberly and Carlin, 2000), but this problem also has some controversies (see, for example, Gelfand and Sahu, 1999) especially using a prior too informative since in some cases the nonidentifiability would not depend on the nature of the prior specification but on lack of identifiability in the likelihood (see also, Kadane, 1975). Kass et al. (1998) point out that provided the posterior is proper, there is no problem for MCMC methods for nonidentifiability problems.

Autores como Kadane (1975) têm estudado as heurísticas (lógicas e empirismos) por trás do processo de eliciação da distribuição a priori. O mesmo Kadane (1975) e Chaloner e Duncan (1983) concentraram-se na construção de métodos de eliciação para diferentes verossimilhanças. Na análise de proporções, geralmente, a especificação da distribuição a priori é baseada numa expressão de probabilidades subjetivas sobre a proporção desconhecida, usando distribuições dentro da família Beta(a,b), a qual contém uma ampla variedade de formas, cujos componentes são conjugados para dados com verossimilhança Binomial, fato que facilita muito a determinação da distribuição a posteriori. Chaloner e Duncan (2001) desenvolveram um algoritmo para obter os hiperparâmetros da distribuição a priori Beta(a,b) usando informação sobre a moda da distribuição preditiva, que é uma Beta-binomial. Esse método foi estudado e aprimorado posteriormente por Gavasakar (1988).

Procedimento proposto:

Definir como θ uma proporção de interesse, uma quantidade aleatória com distribuição de probabilidade desconhecida $\pi(\theta)$. Uma candidata natural para $\pi(\theta)$ é a distribuição Beta com hiperparâmetros a e b . Nesta proposta, consideram-se duas situações distintas para obter os valores dos hiperparâmetros da distribuição Beta(a,b) a priori: quando o analista de dados tem artigos ou publicações de alguma classe com a informação sobre o parâmetro e quando se tem ausência absoluta de informação sobre o parâmetro ou se tem informação tão pouco relevante que a mesma pode ser considerada inútil para elicitar a distribuição a priori.

Procedimento quando se possui informação publicada sobre o parâmetro:

Utilizando-se resultados de relatórios de pesquisa ou artigos, é possível determinar intervalos cujos limites θ_1 e θ_2 contém $(1 - \alpha)\%$ dos valores do parâmetro θ da mesma forma como foi considerado por Joseph et al.(1995). No passo seguinte, iguala-se o valor médio do intervalo (θ_0) com a média da distribuição Beta (a,b), isto é:

$$\theta_0 = \frac{\theta_1 + \theta_2}{2} \cong \frac{a}{a + b}$$

Usar a desigualdade de Chebyshev para obter informação sobre a variância:

$$\begin{aligned}P(|\theta - E(\theta)| \geq k\sigma) &\leq \frac{1}{k^2} = \alpha \\P([\theta - E(\theta)]^2 \geq k^2\sigma^2) &\leq \alpha \\P(\alpha[\theta - \theta_0]^2 \geq \sigma^2) &\leq \alpha\end{aligned}$$

Assim usando θ_1 ou θ_2 ,

$$\sigma^2 \leq \alpha[\theta_1 - \theta_0]^2 \cong \frac{ab}{(a+b)^2(a+b+1)}$$

Procedimento quando não se tem nenhuma informação a priori sobre o parâmetro:

Nas situações em que se desconhece totalmente as características da distribuição de probabilidade a priori, uma alternativa para contornar a situação pode ser:

- (1) Dividir o espaço paramétrico em k intervalos limitados e assumir que o parâmetro está contido entre os limites do intervalo com uma probabilidade igual a $(1 - \alpha)$;
- (2) Com cada um desses intervalos, utilizar o procedimento proposto para obter os valores dos hiperparâmetros;

- (3) Para cada uma das k distribuições Beta(a,b) a priori obtidas, gerar uma amostra de dados simulados usando algum software como o R, observar a forma da distribuição e obter algumas estatísticas de posição e dispersão;
- (4) Com as distribuições a posteriori, realizar uma análise de sensibilidade utilizando-se critérios de seleção de modelos desenvolvidos na literatura (DIC e fator de Bayes, por exemplo), para obter o melhor ajuste aos dados.

Outra possibilidade: usar a distância de Kullback Leibler que permite comparar por pares de distribuições a priori concorrentes.

Exemplos de elicitação de distribuições a priori

(1) Dados de Infecção Urinária

Para este exemplo foram utilizados os dados introduzidos por Ali et al. (2007). Esses autores obtiveram as estimações dos parâmetros de desempenho de dois testes rápidos para triagem de infecção do trato urinário em crianças cujas idades oscilavam entre um mês e onze anos.

O procedimento de diagnóstico incluía como testes de triagem a presença de nitritos ($N=T_1$) e os níveis de esterasa leucocitária na urina ($LE=T_2$) da criança. Como padrão-ouro foi utilizada uma cultura de urina.

Sob o pressuposto de independência entre os resultados dos testes, tem-se um vetor de cinco parâmetros para estimar : duas sensibilidades, duas especificidades e uma prevalência ($\theta = S_j, E_j, p; j = 1, 2$). Os autores compararam os seus resultados usando método de máxima verossimilhança com aqueles obtidos em outros quatro estudos publicados na literatura, tal como aparece na Tabela 6.1.

Tabela 6.1: Resultados de Ali et al. (2007) e de quatro estudos que consideravam os mesmos testes diagnósticos. Médias e variâncias usadas para obter os hiperparâmetros.

Estudo	Incidência %	Presença de nitratos (T_1)		Esterasa leucocítica (T_2)	
		S_1	E_1	S_2	E_2
Ali et al. (2007)	67.4	38.2	88.4	85.4	58.1
Sharief et al. (1998)	5.2	54.6	96.8	100	78.1
Weinberg e Gan. (1991)	4.01	56.0	98.1	85.4	92.7
Lohr et al. (1993)	14.8	37.3	100	79.4	72.7
Cannon et al. (1986)	14.4	72.7	99.6	84.6	71.4
Média θ_0	21.2	51.8	96.6	87.0	74.6
Variância σ^2	963.3	214.4	22.5	59.4	156.5

Usando o procedimento dado acima, temos:

distribuições a priori para os parâmetros de desempenho dos testes N e LE são dadas por: $S_1 \sim Beta(4.15, 4.5)$, $S_2 \sim Beta(15.7, 2.4)$, $E_1 \sim Beta(0.5, 13)$, $E_2 \sim Beta(8.3, 2.8)$, e $p \sim Beta(2.1, 22.3)$.

(2) Dados de câncer de próstata

Para este exemplo foram utilizados os dados do estudo realizado por Smith et al. (1997), que avaliaram a suspeita de câncer de próstata em 19.476 homens adultos maiores

de 50 anos. Os autores usaram como testes de triagem o toque retal (DRE em inglês) e os níveis de antígeno prostático em sangue (PSA em inglês). Níveis de PSA maiores de 4,0 ng/ml foram considerados como suspeitos de câncer. Os indivíduos com resultado positivo em algum dos testes foram verificados com uma biópsia de tecido prostático (padrão-ouro). Neste exemplo, é de interesse estimar o parâmetro de dependência entre DRE e PSA usando uma função cópula de Gumbel.

Nesse estudo, tem-se presença de viés de verificação, de modo que é preciso estimar as quantidades faltantes para construir o arranjo de dados da Tabela 2.2 e definir o modelo de estimação. Utilizando-se as estimativas de δ_k e λ_k (índices de Böhning e Patilea), estimaram-se n_+ e n_- (Tabela 6.2)

Tabela 6.2: Valores estimados para as quantidades de indivíduos não verificados. Os valores entre colchetes são calculados utilizando-se o índice δ e os outros o índice λ

	Indivíduos doentes $\lambda = 2.42, \delta = 3.08$			Indivíduos não doentes $\lambda = 2.40, \delta = 3.03$		
	DRE+	DRE-	Total	DRE+	DRE-	Total
PSA+	189	292	481	141	755	896
PSA-	145	1431[691]	1576[836]	1002	15521[16261]	16523[17263]
Total	334	1723[983]	2057[1317]	1143	16276[17016]	17419[18159]

Como os índices apresentam valores próximos da unidade e similares entre doentes e não doentes, por conseguinte, a dependência entre testes é fraca e positiva nas duas populações.

Utilizando-se os dados da Tabela 6.2 foram obtidas as estimativas preliminares e os intervalos com probabilidade $1 - \alpha = 0,95$ para cada um dos componentes do vetor de parâmetros θ . As estimativas para as variâncias foram calculadas tomando o ponto central do intervalo como θ_0 e aplicando-se o procedimento apresentado na seção 6.1 e o primeiro procedimento proposto por Joseph et al. (1995) descrito nessa mesma seção. Na Tabela 6.3 aparecem os valores calculados para os hiperparâmetros.

Tabela 6.3: Valores dos hiperparâmetros das distribuições a priori para a prevalência, parâmetros dos testes e parâmetro de dependência Gumbel (ϕ) usando o procedimento proposto e o procedimento de Joseph et al. (1995)

θ	<i>Intervalo</i>	θ_0	Método de Joseph		Método proposto	
			<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>
S_1	0.236 - 0.365	0.301	60	140	303	704
S_2	0.162 - 0.254	0.208	65	246	324	1232
E_1	0.949 - 0.951	0.950	180499	9500	902500	47500
E_2	0.934 - 0.937	0.936	100351	6919	501758	34595
p	0.068 - 0.106	0.087	76	800	379	4002
ϕ_1	0.0 - 0.25	0.125	3.4	23.6	17	122
ϕ_2	0.25 - 0.75	0.500	7.5	7.5	39.5	39.5
ϕ_3	0.75 - 1.0	0.875	23.6	3.4	122	17

Examples

Simulated data

We simulated 10,000 pairs of observations with binary dependence structure and the same number of pairs of data for each copula structure (1,000 diseased individuals and 9000 nondiseased individuals), considering the following conditions:

- (1) Three dependence levels: weak (0.2), moderate (0.5) and strong (0.9), in each of the copula structures.
- (2) The specificities of the dependent tests are the same ($E_1 = E_2 = 0.95$) and the prevalence is relatively small ($p = 0.10$)
- (3) The dependent tests have the same relatively high sensitivities ($S_1 = S_2 = 0.85$)

We developed a computer program in the software R to simulate 1,000 pairs of variates with the different dependence structures. To simulate outcomes of variables with FGM structure and with Gumbel structure, we implemented algorithms introduced by Johnson (1987). The Clayton data set was simulated using the approach developed by McNeil et al. (2005) The models were fitted using a Bayesian approach. For weak FGM and Gumbel dependences and the prevalence we have used a Beta(17,122) distribution as informative prior. Moderate dependences given in these copulas were modeled using a

Beta(39.5; 39.5) distribution as informative prior and for strong dependences and test parameters we have used the Beta(122; 17) distribution as informative prior. The models with non-informative prior distributions were fitted using the Beta(1/2; 1/2) prior distribution. The simulated Clayton data set was modeled using the Gamma(20; 10), Gamma(20; 45) and Gamma(20; 1.33) distributions as informative priors on the weak, moderate and strong dependence parameters, respectively. We ran our MCMC algorithm 1,000 times using Winbugs 1.4 software and we obtained the mean and the standard deviations of the Bayesian estimates. For the DIC values, we obtained the mean and the standard deviation from the 1,000 MCMC samples. See Table 4. As we have posterior distributions with not closed forms, we have simulated 500,000 Gibbs samples from the conditional distribution for each parameter. From these generated samples, we discarded the first 50,000 samples to eliminate the effect of the initial values considering a spacing of size 100 to get the final simulated sample. Convergence of the algorithm was verified graphically and also using standard existing methods (Geweke, 1992) implemented in the software CODA (Best et al, 1995). We have used the Deviance Information Criteria (DIC, built within the WinBUGS environment) as proposed by Spiegelhalter et al. (2002) to choose between the proposed models. From DIC criteria, the model with the smallest DIC is better fitted by the data. For the selection model, we also used a heuristic procedure that assumes two criteria: quality in the convergence of the MCMC procedure and concentration of the posterior distribution using the coefficient of variation (CV). The best model should have

the best performance in MCMC convergence and highest concentration around the posterior mean (lowest CV).

Table 4: Means, standard deviations and 95% credibility intervals of the Bayesian estimates to performance test parameters under three copula dependence levels and under independence assumption. Results obtained from 1000 simulated MCMC samples (r = Pearson correlation estimate)

Model	Copula function	Parameter	Weak dependence		Moderate dependence		Strong dependence	
			$\varphi = 0.2$ $\phi = 0.2$ $\alpha = 0.5$	$r = 0.0389$ $r = -0.0504$ $r = -0.0099$	$\varphi = 0.5$ $\phi = 0.5$ $\alpha = 2.0$	$r = 0.0351$ $r = -0.243$ $r = 0.0037$	$\varphi = 0.9$ $\phi = 0.9$ $\alpha = 18$	$r = 0.0633$ $r = -0.3361$ $r = 0.0120$
Models fitted using copula function	FGM	S_1	0.847(0.0105)	0.822-0.871	0.836(0.0136)	0.809-0.862	0.828(0.0161)	0.798-0.855
		S_2	0.851(0.0076)	0.826-0.874	0.834(0.0163)	0.807-0.860	0.819(0.0143)	0.790-0.847
		E_1	0.949(0.0027)	0.944-0.953	0.951(0.0036)	0.946-0.955	0.950(0.0019)	0.945-0.954
		E_2	0.952(0.0018)	0.947-0.956	0.948(0.0020)	0.943-0.952	0.951(0.0021)	0.946-0.955
		P	0.100(0.0005)	0.094-0.106	0.102(0.0007)	0.096-0.108	0.103(0.0007)	0.097-0.109
	DIC		53.25(0.82)		55.83(16.72)		66.38(5.53)	
	GUMBEL	S_1	0.822(0.0080)	0.795-0.847	0.761(0.0197)	0.731-0.790	0.758(0.0141)	0.728-0.787
		S_2	0.806(0.0166)	0.779-0.832	0.766(0.0190)	0.736-0.795	0.762(0.0145)	0.731-0.791
		E_1	0.958(0.0013)	0.954-0.962	0.963(0.0026)	0.959-0.966	0.967(0.0018)	0.964-0.971
		E_2	0.957(0.0022)	0.953-0.961	0.963(0.0019)	0.959-0.967	0.968(0.0021)	0.965-0.972
		P	0.099(0.0007)	0.093-0.105	0.099(0.0023)	0.090-0.101	0.087(0.0010)	0.081-0.105
	DIC		61.15(4.53)		97.08(5.16)		198.44(31.1)	
	CLAYTON	S_1	0.870(0.0105)	0.848-0.891	0.868(0.0140)	0.846-0.889	0.864(0.0155)	0.842-0.885
		S_2	0.866(0.0125)	0.844-0.887	0.867(0.0138)	0.847-0.891	0.864(0.0191)	0.841-0.885
		E_1	0.947(0.0011)	0.943-0.952	0.948(0.0020)	0.944-0.953	0.947(0.0026)	0.943-0.952
E_2		0.946(0.0023)	0.942-0.951	0.947(0.0018)	0.942-0.951	0.948(0.0022)	0.943-0.953	
P		0.098(0.0005)	0.092-0.104	0.098(0.0006)	0.093-0.104	0.098(0.0005)	0.092-0.104	
DIC		69.29(2.97)		69.92(3.08)				
Models fitted under independence assumption	FGM	S_1	0.851(0.0154)	0.832-0.868	0.852(0.0125)	0.835-0.869	0.854(0.0099)	0.818-0.870
		S_2	0.850(0.0062)	0.824-0.866	0.843(0.0131)	0.826-0.860	0.836(0.0099)	0.837-0.853
		E_1	0.946(0.0026)	0.943-0.950	0.949(0.0025)	0.946-0.952	0.950(0.0027)	0.947-0.953
		E_2	0.923(0.0006)	0.919-0.926	0.921(0.0020)	0.917-0.925	0.922(0.0015)	0.918-0.926
		P	0.100(0.0004)	0.096-0.104	0.100(0.0005)	0.096-0.104	0.099(0.0003)	0.096-0.104
	DIC		171.10(8.94)		145.79		162.73(13.4)	
	GUMBEL	S_1	0.815(0.0080)	0.796-0.834	0.738(0.0139)	0.715-0.760	0.705(0.0246)	0.680-0.729
		S_2	0.813(0.0058)	0.794-0.832	0.743(0.0187)	0.720-0.765	0.710(0.0153)	0.685-0.735
		E_1	0.958(0.0014)	0.955-0.961	0.963(0.0027)	0.960-0.966	0.968(0.0052)	0.965-0.970
		E_2	0.926(0.0008)	0.922-0.930	0.930(0.0011)	0.926-0.934	0.935(0.0025)	0.931-0.939
		P	0.099(0.0006)	0.095-0.103	0.098(0.0017)	0.094-0.102	0.095(0.0020)	0.090-0.099
	DIC		189.28(9.44)		234.57(9.17)		268.81(19.86)	
	CLAYTON	S_1	0.853(0.0144)	0.836-0.869	0.850(0.0139)	0.832-0.866	0.843(0.0198)	0.825-0.859
		S_2	0.844(0.0065)	0.826-0.861	0.855(0.0192)	0.838-0.871	0.837(0.0190)	0.819-0.854
		E_1	0.951(0.0011)	0.948-0.954	0.951(0.0025)	0.948-0.954	0.948(0.0016)	0.945-0.951
E_2		0.921(0.0005)	0.917-0.924	0.921(0.0007)	0.917-0.925	0.922(0.0014)	0.918-0.926	
P		0.100(0.0007)	0.096-0.105	0.100(0.0008)	0.096-0.104	0.100(0.0002)	0.096-0.105	
DIC		151.32(8.88)		156.95(14.36)		163.62(13.95)		

In agreement with the results showed in Table 4, the linear correlations associated with the assumed copula dependences are very weak taking values that can easily lead to the conclusion that the tests are independent. The

standard deviations obtained for all sensitivity MCMC estimates are higher than those observed to other parameter estimates even where the convergence of the MCMC process in each run is good. The three dependence types affect the sensitivity estimates; the "copula1-model" and "copula2-model" tend to underestimate and the "copula3-model" shows an opposite behaviour. When we fitted a model assuming independence between test outcomes, we observed that for high dependences type "copula1-model" and "copula3-model" a sensitivity estimate tends to be underestimated while, for "copula2-model" both sensitivity estimates are underestimated and the estimation bias is higher than that observed in the models that consider the dependence. In all models fitted under independence assumption, one of the specificities is underestimated. (see Table 4).

Cancer Data

To illustrate the method described in the previous sections, we use the data from the Health Insurance Plan Study (HIP) for breast cancer screening in New York, as described by Strax et al. (1967) and previously analyzed by Schatzkin et al. (1987), Cheng and Macaluso (1996), Walter (1999) and Van der Merwe and Maritz (2002). In this study, 20,211 women were evaluated by mammography and physical examination as screening tests to detect breast cancer. The disease status was determined by biopsy. The mammography and physical examination are not basically biological traits with outcomes in a continuous scale. In this case, the test results depend on knowledge and experience of the

radiologist analysing the images and evaluating the clinical symptoms. Therefore doctor's knowledge in both areas can be assumed as two non-observable traits present in the same individual that can be measured building a continuous scale in the same way that is done with educational and psychological traits. The data of the study is given in Table 5.

Table 5: Data set (Breast cancer data)

Physical Examination	Positive Biopsy		Negative Biopsy	
	Mammography +	Mammography -	Mammography +	Mammography -
+	10	24	13	144
-	21		95	

Considering the results given by Walter (1999), we built 95% confidence intervals and from these intervals we applied the method described by Joseph et al. (1995) to obtain the hyperparameters for the Beta informative prior distributions used on prevalence and test parameters. (See Table 6).

Table 6: Prior distribution hyperparameters for performance test parameters

Parameter	Interval	$E(\theta)$	a_θ	b_θ
$\theta_{1,1} = S_1$	0.1516 - 0.4872	0.3216	9.7	20.3
$\theta_{1,2} = S_2$	0.1391 - 0.4473	0.2932	9.9	24.0
$\theta_{1,3} = E_1$	0.9910 - 0.9934	0.9922	21329	168
$\theta_{1,4} = E_2$	0.9936 - 0.9956	0.9946	21366	116
$\theta_{1,5} = p$	0.0027 - 0.0078	0.0052	16.9	3228

For "copula1-model" and "copula2-model" dependence parameters, we do not have any prior information on the form how it is expressed the dependence measured by the copulas in the data set. In this way, we decided to divide the parametric space in three intervals of values where it could be possible to find the dependence parameter, namely; $(0, 1/4)$, $(1/4, 3/4)$ and $(3/4, 1)$. Using these three intervals of values, we have used the Chebychev inequality to approximate the prior means and variances for the dependence parameters. We assumed Beta(17; 122), Beta(39.5; 39.5) and Beta(122; 17) distributions as informative priors for the "copula1-model" and "copula2-model" dependences, respectively.

For the "copula3-model" dependence parameter, we have used the relationship between α and the Kendall tau (equation (13)) also assuming that ζ belongs to each interval built for the other two copula models. Thus, for the α parameter, we have three intervals as follows:

$$\alpha \in (0, 2/3) \text{ when } \tau \in (0, 1/4), \quad \alpha \in (2/3, 6) \text{ when } \tau \in (1/4, 3/4) \text{ and } \alpha \in (6, 98) \text{ if } \tau \in (3/4, 0.99).$$

With these intervals we obtained three informative prior distributions namely: $\text{Gamma}(20, 45)$,

$\text{Gamma}(20, 7.5)$ and $\text{Gamma}(20, 1.333)$

As non-informative prior distributions, we have used a $U(0,1)$ distribution for "copula 1-model" and "copula 2-model" dependences and $Beta(1/2, 1/2)$ distribution for prevalence and performance test parameters. To elicit a non-informative prior distribution for the α parameter, we assume that $\alpha \in (0, 1000)$, and from the Chebychev inequality we obtained the hyperparameters $r = 20$ and $s = 0.04$ for the gamma prior distribution.

For each dependence copula parameter, we fitted seven models, one with informative prior on performance tests and non-informative prior for the dependence parameter, three with non-informative prior over test parameters and informative prior over dependence parameters and finally, three models with informative prior distributions for all parameters. Using our heuristic criteria, we selected six models, two for each dependence coefficient and we divided them in two groups, one with those models using informative prior distributions for dependence parameters and non-informative prior distributions for test parameters (Set 1) and one group with models that have prior informative for all parameters (Set 2). We also fitted two models under independence assumption, one of them assuming $Beta(1/2; 1/2)$ as non-informative prior distribution for each performance test parameter (first model in Set 1) and the other one with $Beta(a; b)$ as informative prior distributions using the values showed in Table 6 as hyperparameters (Set 2). The results are given in Table 7.

Using the results obtained by Walter (1999), we calculated an approximation of the Pearson correlation coefficient in both populations ($r_D = 0.00403$ and $r_{ND} = 0.09327$), whose values could be indicating that the tests have very weak positive linear dependence. The results obtained with the model fitted under independence assumption are very similar with those obtained by Walter (1999) using the maximum likelihood approach.

Table 7: Bayesian results assuming independence between tests and three copula dependence structures. (Posterior means and 95% credible intervals (95% ICr) for each parameter of interest)

Set 1 of models				Set 2 of models			
Model	Parameter	Means	95%ICr	Model	Parameter	Means	95%ICr
$M_{1,1}$ $DIC = 87.62$	S_1	0.3229	0.2137 - 0.4415	$M_{2,1}$ $DIC = 100.59$	S_1	0.3246	0.2444 - 0.4130
	S_2	0.2945	0.1927 - 0.4077		S_2	0.2954	0.2206 - 0.3782
	E_1	0.9922	0.9913 - 0.9930		E_1	0.9922	0.9915 - 0.9929
	E_2	0.9946	0.9932 - 0.9957		E_2	0.9963	0.9955 - 0.9969
	p	0.0054	0.0039 - 0.0076		p	0.0052	0.0041 - 0.0066
$M_{1,2}$ $DIC = 75.29$	φ_D	0.8774	0.8180 - 0.9264	$M_{2,2}$ $DIC = 74.86$	φ_D	0.8749	0.8150 - 0.9245
	φ_{ND}	0.8822	0.8252 - 0.9294		φ_{ND}	0.8824	0.8257 - 0.9297
	S_1	0.2127	0.1047 - 0.3548		S_1	0.3320	0.2399 - 0.4337
	S_2	0.1941	0.0944 - 0.3272		S_2	0.3099	0.2218 - 0.4092
	E_1	0.9921	0.9909 - 0.9933		E_1	0.9922	0.9913 - 0.9930
	E_2	0.9946	0.9935 - 0.9955		E_2	0.9946	0.9939 - 0.9953
$M_{1,3}$ $DIC = 114.94$	p	0.0086	0.0049 - 0.0153	$M_{2,3}$ $DIC = 110.89$	p	0.0054	0.0041 - 0.0071
	ϕ_D	0.1216	0.0732 - 0.1803		ϕ_D	0.1227	0.0739 - 0.1813
	ϕ_{ND}	0.0380	0.0222 - 0.0580		ϕ_{ND}	0.0379	0.0222 - 0.0576
	S_1	0.3698	0.2257 - 0.5292		S_1	0.3806	0.2799 - 0.4913
	S_2	0.3374	0.2033 - 0.4891		S_2	0.3570	0.2593 - 0.4651
	E_1	0.9921	0.9908 - 0.9932		E_1	0.9921	0.9913 - 0.9929
$M_{1,4}$ $DIC = 141.54$	E_2	0.9945	0.9934 - 0.9955	$M_{2,4}$ $DIC = 115.81$	E_2	0.9946	0.9938 - 0.9952
	p	0.0047	0.0032 - 0.0071		p	0.0046	0.0035 - 0.006
	α_D	0.3308	0.2071 - 0.4774		α_D	0.4400	0.2703 - 0.6518
	α_{ND}	0.4415	0.2693 - 0.6572		α_{ND}	1.310^{-5}	$8.910^{-7} - 1.810^{-6}$
	S_1	0.0017	0.0012 - 0.0024		S_1	0.3927	0.2872 - 0.5073
	S_2	0.0016	0.0011 - 0.0022		S_2	0.3687	0.2664 - 0.4800
	E_1	0.5556	0.5032 - 0.6074	E_1	0.9910	0.9896 - 0.9922	
	E_2	0.6932	0.6466 - 0.7375	E_2	0.9940	0.9933 - 0.9948	
	p	0.9824	0.9800 - 0.9847	p	0.0044	0.0033 - 0.0058	

$M_{j,1}$, $j = 1, 2$: Models under assumption of independence between tests

$M_{j,2}$, $j = 1, 2$: "copula1-model" dependence parameters with $Beta(122, 17)$ prior distributions

$M_{j,3}$, $j = 1, 2$: "copula2-model" dependence parameters with $Beta(17, 122)$ prior distributions

$M_{j,4}$, $j = 1, 2$: "copula3-model" dependence parameters with $Gamma(20, 45)$ prior distributions

In agreement with the observed results, the best fit is obtained with the model that uses "copula1-model" dependence, which shows 95% credibility intervals shorter than 95% confidence intervals obtained with the Walter (1999) results and sensitivity estimators a little higher. This model showed the lowest DIC among all evaluated models. It is important to point out that, the models with informative prior distributions on all parameters have DIC values lower than those obtained with models fitted assuming informative prior distributions on copula parameter and non-informative prior on the other parameters; this result was not observed only in the model fitted under independence assumption.

Conclusions and remarks

We have used copula functions to study the dependence between two diagnostic tests selecting six models between 21 evaluated models considering as selection criterion, the DIC, the posterior coefficient of variation and MCMC convergence performance. We considered the diagnostic designs under verification bias that use two diagnostic tests with continuous outcome applied on the same individual and dichotomized to develop the data analysis.

In those cases, given that the observations are binary, many authors have used binary covariance to study the dependence between test results but this approach does not consider the original continuous structure in the data.

The first copula function models data with very weak linear dependence (FGM copula), the second one models data with weak negative dependences (Gumbel) and the third one it is useful when the dependence is mainly concentrated in the lower tail (Clayton copula).

The developed reparametrization allows us to obtain the prevalence and test parameter estimates using the binary data inside a model that includes the structure of dependence of the continuous data and where it is possible to obtain the parameter estimates under a Bayesian approach in a very easy way using a computer program written in Winbugs 1.4 software.

Data of diagnostic tests with very weak linear dependences apparently could be well fitted by a model assuming independence, but as a result of our simulation study, we observe that if the data have one of the copula dependence structures studied here, at least one of the test parameters will be underestimated and in cases as the "copula2-dependence" three of the five estimates show a estimation bias higher than that observed when we fit the data with a model that considers dependence type copula.

We illustrated the proposed methodology using a published data set and we obtained estimation intervals shorter than those obtained using other methodologies as the latent variables under maximum likelihood approach and generalized linear models. The model fitted using the Clayton copula showed the worst results and convergence problems for the Gibbs sampling algorithm. This could be due to the fact that this copula is only

appropriated to model dependences concentrated in the tails, a fact not observed in the used data set.

References

ACHCAR, J. A.; MARTINEZ, E.; LOUZADA N. F.. Binary data in the presence of covariates and misclassifications: A Bayesian approach. **Brazilian Journal of Probability and Statistics**. v.19, p.65-84, 2005.

ALBERT, Paul; DODD, Lori. A cautionary note on the robustness of latent class models for estimating diagnostic error without a gold standard. **Biometrics**. v.60, p.427-435, 2001.

ALI, S.; MOODAMBAIL, A.R.; HAMRAH, E.; BINNAKHI, H.; SADEQ, S.. Reliability of rapid dipstick test in detecting urinary tract infection in symptomatic children. **Kuwait Medical Journal**. v.39, p.36-38, 2007.

BAKER, S.Evaluating multiple diagnostic tests with partial verification. **Biometrics**. v.51, p.330-337, 1995.

BEGG, C.. Biases in the assessment of diagnostic tests. **Statistics in Medicine**. v.6, p.411-423, 1987.

BLACKSTONE, E.; LAUER, M.Caveat emptor: The treachery of work-up bias. **The Journal of Thoracic and Cardiovascular Surgery** v.128, n.3. p.341-344, 2004.

BÖHNING, D.; PATILEA, V.. A capture-recapture approach for screening using two diagnostic tests with availability of disease status for the test positives only. **Journal of American Statistical Association**. v.103, n.481, p.212-221, march, 2008.

CHALONER, K., DUNCAN, G. Assessment of a Beta prior distribution: PM Elicitation. **The Statistician**. v.32, p.174-180, 1983.

CHENG, H.; MACALUSO M.. Comparison of the accuracy of two tests with a confirmatory procedure limited to positive results. **Epidemiology**. v.8, p.104-106, 1997.

CLAYTON, D.G. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. **Biometrika**. v.65, p.141-151, 1978.

DAWID, A. P. Conditional independence in statistical theory (with discussion). **Journal of the Royal Statistical Society**, Series B 41, 1–31, 1979

DENDUKURI, N.; JOSEPH, L.. Bayesian approach to modeling the conditional dependence between multiple diagnostic tests. **Biometrics**. v.57, p.158-167, 2001.

EBERLY, L. and B. CARLIN . Identifiability and convergence issues for Markov Chain Monte Carlo fitting of spatial models. **Statistics in Medicine** 19, 2279–2294, 2000.

GARRET, E.; EATON, W.; ZEGER, S.. Methods for evaluating the performance of diagnostic tests in the absence of a gold standard: a latent class model approach. **Statistics in Medicine**, v.21, p.1289-1307, 2002.

GAVASAKAR, U.. A comparison of two elicitation methods for a prior distribution for a Binomial parameter. **Management science**. v.34, n.6, p.784-790, June 1988.

GELFAND, A. and SAHU ,S. Identifiability, Improper Priors, and Gibbs Sampling for Generalized Linear Models. **Journal of the American Statistical Association** ,94, 247–253, 1999.

GUPTA, A.; ROEHRBORN, C. Verification and incorporation biases in studies assessing screening tests: prostate-specific antigen as an example. **Urology**. v.64, n.1, p.106-116, July, 2004.

HUI, .S.; ZHOU, X. Evaluation of diagnostic tests without gold - standard. **Statistical Methods in Medical Research**. v.7, p.354-370, 1998.

JOHNSON, M.. **Multivariate Statistical Simulation**. John Wiley and Sons, 1987.

JOSEPH, L.; GYORKOS, T.; COUPAL, L.. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. **American Journal of Epidemiology**. v.141, p.263-272, 1995.

KADANE, J. The role of identification in bayesian theory. In S. Fienberg and A. Zellner (Eds.), **Studies in Bayesian Econometrics and Statistics**, pp. 175–191. Amsterdam: North-Holland,1975

KASS, R., B. CARLIN, A. GELMAN, and R. NEAL . Markov Chain Monte Carlo in Practice: A Roundtable Discussion. **The American Statistician**,52, 93–100,1998.

LINDLEY, D. V. Bayesin Statistics: A Review. Philadelphia: **Society for Industrial and Applied Mathematics**,1971

MARTINEZ, E.; ACHCAR, J.; LOUZADA NETO, F.. Análise Bayesiana do desempenho de dois testes diagnósticos quando indivíduos com resultados negativos em ambos os testes não são verificados por um padrão-ouro. **Revista Brasileira de Biometria**. v.22, n.3, p.21-32, 2004

MARTINEZ, E.; ACHCAR, J.; LOUZADA NETO, F.. Bayesian estimation of diagnostic tests accuracy for semi-latent data with covariates. **Journal of Biopharmaceutical Statistics**. v.15, p.809-821,2005.

MARTINEZ, E.; ACHCAR, J.; LOUZADA NETO, F.. Estimatorsof sensitivity and specificity in the presence of verification bias: A Bayesian approach. **Computational Statistics and Data Analysis**. v.51, p.601-611, 2006.

MARTINEZ, E.; LOUZADA NETO, F.; DERCHAIN, S. F.;ACHCAR, J.; GONTIJO R.SARIAN, L.; SYRJÄNEN, K..Bayesian estimation of performance measures of cervical cancer screening tests in the presence of covariates and absence of a gold standard. **Cancer Informatics**. v.6, p.33-46, 2008.

MARTINEZ, E.; LOUZADA NETO, F.; ACHCAR, J.; SYRJÄNEN,K.; DERCHAIN, S.F.; GONTIJO R.; SARIAN, L.. Bayesian estimation of performance measures of screening tests in the presence of covariates and absence of a gold standard. **Brazilian Journal of Probability and Statistics**. v.23, p.68-81, 2009.

McNEIL, A.; FREY, R.; EMBRECHTS, P.. **Quantitative Risk Management: Concepts, Techniques and Tools**. Pricenton University Press, 2005 apud SCHMITDT T.. Coping with Copulas. In: RANK J.(org.). **Copulas: From theory to application in finance**.

Edited by Rank J. Director of d-fine, one of Europe's leading financial service consulting firms,2005.

PAULINO, C D. and C. PEREIRA, C.A.. On identifiability of parametric models. **Journal of the Italian Statistical Society** 3, 125–151,1994

PEPE, M.S., ALONSO, T.A. Comparing disease screening tests when true disease status is ascertained only for screen positives. **Biostatistics**. 2:249-260,2001.

POIRIER, D. J. Revising Beliefs in Nonidentified Models. **Econometric Theory** 14, 483–509,1998.

QU, Y.;MING, T; KUTNER, M. Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. **Biometrics**, v.52, p.797-810, 1996.

QU, Y.; HADGU, A.. A model for evaluating sensitivity and specificity for correlated diagnostic test in efficacy studies with an imperfect reference test. **Journal of the American Statistical Association**. v.93, n.443, p.920-928, 1998.

SCHATZKIN, A.; CONNOR, R.; TAYLOR, P.; BUNNAG, B. Comparing new and old screening tests when a reference procedure cannot be performed on all screenees. **American Journal of Epidemiology**. v.125, n.4, p.672-678, 1987.

SKLAR, A. Fonctions de répartition à n dimensions et leurs margens. **Publ. Inst. Statist. Univ. Paris**. v.8, p.229-231, 1959, apud NELSEN, Roger. **An introduction to copulas**. Springer Verlag, New York, 1999.

SMITH, D.; BULLOCKA, A.; CATALONA, W.. Racial differences in operating characteristics of prostate cancer screening tests. **Journal of Urology**. V.158, n.5, p.1861-1865, 1997.

SPIEGELHALTER, D.; BEST, N.; CARLIN, B.; VAN DER LINDE, A.. Bayesian measures of model complexity and fit. **Journal of the Royal Statistical Society, Series B**. v.64, n.4, p.583-639, 2002.

STAQUET M.; ROZENCWEIG, M.; LEE, Y.; MUGGIA, F. Methodology for the assessment of new dichotomous diagnostic tests. **Journal of Chronical Diseases**. v.34, p. 599-610, 1981.

STRAX, P., VENER, L., SHAPIRO, S., GROSS, S. (1967). Mammography and Clinical Examination in mass screening for cancer of breast. **Cancer**. 20:2184-21-88, 1967.

TANNER, M.A., WONG, W.H. The calculation of posterior distributions by data augmentation. **Journal of the American Statistical Association**. 82:528-540, 1987

THIBODEAU, L.A. Evaluating diagnostic tests. **Biometrics**. v.37, p.801-804, 1981.

TORRANCE-RYNARD, V.; WALTER, S. Effects of dependent -errors in the assessment of diagnostic tests performance. **Statistics in Medicine**. v.16, p.2157-2175, 1997.

VACEK, P. The effect of conditional dependence on the evaluation of diagnostic tests. **Biometrics**. v.41, p.959-968, 1985.

WALTER, S. Estimation of test sensitivity and specificity when disease confirmation is limited to positive results. **Epidemiology**. v.10, p.67-72, 1999.

VAN der MERWE, L., MARITZ, J.S. Estimating the conditional false-positive rate for semi-latent data. **Epidemiology**. 13:424-430, 2002.

WHITING, P.; RUTJES, A.; REITSMA, J.; GLAS, A.; BOSSUYT, P.; KLEIJNEN, J.. Sources of variation and bias in studies of diagnostic accuracy. A systematic review. **Annals of Internal Medicine**. v.140, p.189-202, 2004.

YANG, I.; BECKER, M. Latent variable modeling of diagnostic accuracy. **Biometrics**. v.53, p.948-958, 1997.

ZHOU, X.. Correcting for verification bias in studies of a diagnostic test's accuracy. **Statistical Methods in Medical Research**. v.7, p.337-353, 1998.