Biometrika (2015), **0**, 0, *pp*. 1–7

Printed in Great Britain

Supplementary material for Quantile-based classifiers

BY C. HENNIG

Department of Statistical Science, University College London, London, WC1E 6BT, U.K.
c.hennig@ucl.ac.uk

AND C. VIROLI

Department of Statistical Sciences, University of Bologna, Bologna, 40126, Italy
cinzia.viroli@unibo.it

1. THEORY FOR $p = 1$

Assuming the notation of Section 2 of the paper, here are some results for $p = 1$.

The following lemma provides a useful formula to derive the theoretical rate of correct classification as function of θ for $p = 1$. This was used for producing Figure 1 of the paper.

LEMMA 1. *When $p = 1$, the probability of correct classification of the quantile classifier takes the following simple form:*

- *if $q_0(\theta) \leq q_1(\theta)$,*

$$\Psi(\theta) = \pi_0 F_0(\ddot{q}_\theta) + \pi_1\{1 - F_1(\ddot{q}_\theta)\} \quad (1)$$

with $\ddot{q}_\theta = \theta q_0(\theta) + (1-\theta) q_1(\theta)$;

- *if $q_0(\theta) > q_1(\theta)$,*

$$\Psi(\theta) = \pi_1 F_1(\dot{q}_\theta) + \pi_0\{1 - F_0(\dot{q}_\theta)\} \quad (2)$$

with $\dot{q}_\theta = \theta q_1(\theta) + (1-\theta) q_0(\theta)$;

where $q_0(\theta)$ and $q_1(\theta)$ are the true quantiles of the two populations.

Proof of Lemma 1.

In the univariate case $\Phi_0(z,\theta)$ and $\Phi_1(z,\theta)$ may be rewritten as

$$\Phi_0(z,\theta) = (1-\theta)\{q_0(\theta) - z\}\mathbb{1}_{\{z \leq q_0(\theta)\}} + \theta\{z - q_0(\theta)\}\mathbb{1}_{\{z > q_0(\theta)\}},$$
$$\Phi_1(z,\theta) = (1-\theta)\{q_1(\theta) - z\}\mathbb{1}_{\{z \leq q_1(\theta)\}} + \theta\{z - q_1(\theta)\}\mathbb{1}_{\{z > q_1(\theta)\}}.$$

For a fixed θ, the integral (5) in the paper can be simplified by splitting it into four parts according to the possible disjoint regions of the domain of Z with respect to $q_0(\theta)$ and $q_1(\theta)$, namely: (a) $z \leq \min\{q_0(\theta), q_1(\theta)\}$, (b) $q_0(\theta) < z \leq q_1(\theta)$, (c) $q_1(\theta) \leq z \leq q_0(\theta)$ and (d) $z > \max\{q_0(\theta), q_1(\theta)\}$.

If $z \leq \min\{q_0(\theta), q_1(\theta)\}$, case (a), the integral becomes

$$
\begin{aligned}
\Psi_a(\theta) = \pi_0 \int_{-\infty}^{\min\{q_0(\theta),q_1(\theta)\}} & \mathbb{1}_{[\{1-\theta\}\{q_1(\theta)-q_0(\theta)\}>0]} dP_0(z) \\
+\pi_1 \int_{-\infty}^{\min\{q_0(\theta),q_1(\theta)\}} & \mathbb{1}_{[\{1-\theta\}\{q_1(\theta)-q_0(\theta)\}\leq 0]} dP_1(z) \\
= \pi_0 \int_{-\infty}^{q_0(\theta)} & dP_0(z) \mathbb{1}_{\{q_1(\theta)>q_0(\theta)\}} + \pi_1 \int_{-\infty}^{q_1(\theta)} dP_1(z) \mathbb{1}_{\{q_1(\theta)\leq q_0(\theta)\}} \\
= \pi_0 \theta & \mathbb{1}_{\{q_1(\theta)>q_0(\theta)\}} + \pi_1 \theta \mathbb{1}_{\{q_1(\theta)\leq q_0(\theta)\}}.
\end{aligned}
$$

In case (b) the integral is

$$
\begin{aligned}
\Psi_b(\theta) = \pi_0 & \int_{q_0(\theta)}^{q_1(\theta)} \mathbb{1}_{[\{1-\theta\}\{q_1(\theta)-z\}-\theta\{z-q_0(\theta)\}>0]} dP_0(z) \\
& +\pi_1 \int_{q_0(\theta)}^{q_1(\theta)} \mathbb{1}_{[\{1-\theta\}\{q_1(\theta)-z\}-\theta\{z-q_0(\theta)\}\leq 0]} dP_1(z) \\
& = \pi_0 \int_{q_0(\theta)}^{\theta q_0(\theta)+(1-\theta)q_1(\theta)} dP_0(z) \mathbb{1}_{\{q_0(\theta)\leq q_1(\theta)\}} \\
& +\pi_1 \int_{\theta q_0(\theta)+(1-\theta)q_1(\theta)}^{q_1(\theta)} dP_1(z) \mathbb{1}_{\{q_0(\theta)\leq q_1(\theta)\}}.
\end{aligned}
$$

Similarly, for cases (c) and (d) the integrals are

$$
\begin{aligned}
\Psi_c(\theta) = \pi_0 & \int_{\theta q_1(\theta)+(1-\theta)q_0(\theta)}^{q_0(\theta)} dP_0(z) \mathbb{1}_{\{q_1(\theta)\leq q_0(\theta)\}} \\
& +\pi_1 \int_{q_1(\theta)}^{\theta q_1(\theta)+(1-\theta)q_0(\theta)} dP_1(z) \mathbb{1}_{\{q_1(\theta)\leq q_0(\theta)\}},
\end{aligned}
$$

and

$$
\Psi_d(\theta) = \pi_0(1-\theta) \mathbb{1}_{\{q_0(\theta)>q_1(\theta)\}} + \pi_1(1-\theta) \mathbb{1}_{\{q_0(\theta)\leq q_1(\theta)\}}.
$$

When $q_0(\theta) \leq q_1(\theta)$, $\Psi(\theta)$ is the sum of $\Psi_a(\theta)$, $\Psi_b(\theta)$ and $\Psi_d(\theta)$ corresponding to disjoint domain regions of Z:

$$
\begin{aligned}
\Psi(\theta) &= \pi_0\theta + \pi_0 \int_{q_0(\theta)}^{\theta q_0(\theta)+(1-\theta)q_1(\theta)} dP_0(z) + \pi_1 \int_{\theta q_0(\theta)+(1-\theta)q_1(\theta)}^{q_1(\theta)} dP_1(z) + \pi_1(1-\theta) \\
&= \pi_0\theta + \pi_0 F_0\{\theta q_0(\theta) + (1-\theta)q_1(\theta)\} - \pi_0\theta + \pi_1\theta - \pi_1 F_1\{\theta q_0(\theta) + \\
&\quad +(1-\theta)q_1(\theta)\} + \pi_1(1-\theta) = \pi_0 F_0(\ddot{q}_\theta) + \pi_1\{1 - F_1(\ddot{q}_\theta)\}.
\end{aligned}
$$

Analogously, when $q_0(\theta) > q_1(\theta)$, $\Psi(\theta)$ is the sum of $\Psi_a(\theta)$, $\Psi_c(\theta)$ and $\Psi_d(\theta)$, from which

$$
\begin{aligned}
\Psi(\theta) &= \pi_1\theta + \pi_0 \int_{\theta q_1(\theta)+(1-\theta)q_0(\theta)}^{q_0(\theta)} dP_0(z) + \pi_1 \int_{q_1(\theta)}^{\theta q_1(\theta)+(1-\theta)q_0(\theta)} dP_1(z) + \pi_0(1-\theta) \\
&= \pi_1 F_1(\dot{q}_\theta) + \pi_0\{1 - F_0(\dot{q}_\theta)\}.
\end{aligned}
$$

Lemma 1 can be reformulated by relabeling the quantities involved in (1) and (2) as follows.

COROLLARY 1. *Let* $q^{(0)}(\theta) = \min\{q_0(\theta), q_1(\theta)\}$ *and* $q^{(1)}(\theta) = \max\{q_0(\theta), q_1(\theta)\}$. *The probability of correct classification is*

$$\Psi(\theta) = \pi^{(0)} F^{(0)}(\tilde{q}_\theta) + \pi^{(1)}\{1 - F^{(1)}(\tilde{q}_\theta)\}, \tag{3}$$

where $\tilde{q}_\theta = \theta q^{(0)}(\theta) + (1-\theta) q^{(1)}(\theta)$. *The other quantities have been renamed accordingly by introducing the notation* $l_\theta = \arg\min_{k=0,1}\{q_k(\theta)\}$ *and* $u_\theta = \arg\max_{k=0,1}\{q_k(\theta)\}$, *from which* $F^{(0)} = F_{l_\theta}$, $F^{(1)} = F_{u_\theta}$, $\pi^{(0)} = \pi_{l_\theta}$ *and* $\pi^{(1)} = \pi_{u_\theta}$.

Analogously, the theoretical misclassification rate of the quantile classifier for $p = 1$ *is*

$$1 - \Psi(\theta) = \pi^{(0)} \left\{1 - F^{(0)}(\tilde{q}_\theta)\right\} + \pi^{(1)} F^{(1)}(\tilde{q}_\theta), \tag{4}$$

with $\tilde{q}_\theta = \theta q^{(0)}(\theta) + (1-\theta) q^{(1)}(\theta)$.

LEMMA 2. *Assume that the cumulative distribution functions of the two populations are such that the density functions* $f^{(0)}(z) = F'^{(0)}(z)$ *and* $f^{(1)}(z) = F'^{(1)}(z)$ *exist for* z *and are nonzero on the same domain. Further assume that there is a point* z_0 *with* $\pi^{(0)} f^{(0)}(z_0) = \pi^{(1)} f^{(1)}(z_0)$ *so that* $\pi^{(0)} f^{(0)}(z) > \pi^{(1)} f^{(1)}(z)$ *for* z *on one side of* z_0 *and* $\pi^{(0)} f^{(0)}(z) < \pi^{(1)} f^{(1)}(z)$ *for* z *on the other side of* z_0. *Then the quantile classifier using the quantile* $\tilde{q}_\theta$ *that minimizes the theoretical misclassification probability achieves the optimal Bayes misclassification probability.*

Proof of Lemma 2. The optimal value θ that minimizes the theoretical misclassification probability in Corollary 1 can be obtained by setting the first derivative of (4) to zero, from which

$$\pi^{(0)} f^{(0)}(\tilde{q}_\theta) = \pi^{(1)} f^{(1)}(\tilde{q}_\theta).$$

A value θ exists so that indeed $\tilde{q}_\theta = z_0$ fulfills this, because under the given assumptions, $q^{(0)}$ and $q^{(1)}$ are continuous functions of θ that converge to the lower and upper bounds of the domain for $\theta \to 0$ and $\theta \to 1$. Furthermore, under the assumptions of the Theorem, the optimal Bayesian classifier only has a single decision boundary at z_0. This proves the Lemma.

2. ASYMPTOTIC THEORY

2·1. *Theorem 1*

This section contains the proofs of Theorems 1 and 2 of the paper, using the notation of Section 3. The assumptions are:

A1. For all $j = 1, \ldots, p,\ k = 0, 1,\ q_{kj}$ is a continuous function of $\theta \in T$.
A2. For all $\theta \in T$, pr $\left[\sum_{j=1}^{p}\{\Phi_{1j}(Z,\theta) - \Phi_{0j}(Z,\theta)\} = 0\right] = 0$.

THEOREM 1. *Under A1 and A2, for any* $\epsilon > 0$,

$$\lim_{n\to\infty} \text{pr}\{|\Psi(\tilde{\theta}) - \Psi(\hat{\theta}_n)| > \epsilon\} = 0.$$

Proof of Theorem 1. The inequality

$$|\Phi_j(z, \theta_1, q_1) - \Phi_j(z, \theta_2, q_2)| \le |z_j||\theta_2 - \theta_1| + 4|q_2 - q_1| \tag{5}$$

is proven below for $j = 1, \ldots, p$ as Lemma 3. Together with A1, this implies the continuity of Ψ, because for given z, Φ_{kj} is a continuous function of θ, and the dominated convergence theorem makes the integrals of the indicator functions converge for $\hat{\theta}_n \to \theta$.

The proof of Theorem 1 is now based on the inequality

$$|\Psi(\tilde{\theta}) - \Psi(\hat{\theta}_n)| \leq |\Psi(\tilde{\theta}) - \Psi_n(\tilde{\theta})| + |\Psi_n(\tilde{\theta}) - \Psi_n(\hat{\theta}_n)| + |\Psi_n(\hat{\theta}_n) - \Psi(\hat{\theta}_n)|. \quad (6)$$

In order to show that all three terms on the right side are asymptotically small, the following result is proved below as Lemma 4.

$$\text{For all } \epsilon > 0 \; \lim_{n\to\infty} \text{pr}\left\{\sup_{\theta \in T} |\Psi_n(\theta) - \Psi(\theta)| > \epsilon\right\} = 0. \quad (7)$$

Result (7) forces the first and third term on the right side of (6) to converge to zero in probability. Now consider the second term. By definition,

$$\Psi_n(\hat{\theta}_n) \geq \Psi_n(\tilde{\theta}), \; \Psi(\tilde{\theta}) \geq \Psi(\hat{\theta}_n).$$

Using (7) again, for large n both $|\Psi_n(\tilde{\theta}) - \Psi(\tilde{\theta})|$ and $|\Psi_n(\hat{\theta}_n) - \Psi(\hat{\theta}_n)|$ will be arbitrarily small with probablity arbitrarily close to 1, and this makes $|\Psi_n(\tilde{\theta}) - \Psi_n(\hat{\theta}_n)|$ arbitrarily small, too.

LEMMA 3. *Equation (5) holds for* $j \in \{1, \ldots, p\}, \theta_1, \theta_2 \in (0,1), \; q_1, q_2 \in \mathbb{R}$ *under the assumption that (a)* $\theta_1 \leq \theta_2$ *and* $q_1 \leq q_2$ *or (b)* $\theta_1 \geq \theta_2$ *and* $q_1 \geq q_2$.

Remark. If both q_i (i=1,2) are θ_i-quantiles of the same distribution, one of (a) or (b) holds.

Proof of Lemma 3: Assume (a) $q_1 \leq q_2, \; 0 < \theta_1 \leq \theta_2 < 1$; case (b) works by analogy. Consider $z_j \leq q_1, \; q_1 < z_j < q_2, \; q_2 \leq z_j$ separately; first $z_j \leq q_1$. By definition,

$$\begin{aligned} |\Phi_j(z, \theta_1, q_1) - \Phi_j(z, \theta_2, q_2)| &= |(1-\theta_1)(q_1 - z_j) - (1-\theta_2)(q_2 - z_j)| \\ &= |(q_1 - q_2) + (\theta_1 + \theta_2)(q_2 - q_1) - \theta_1 q_2 + \theta_2 q_1 + z_j(\theta_1 - \theta_2)| \\ &\leq |q_2 - q_1| + |\theta_1 + \theta_2||q_2 - q_1| + \theta_2|q_2 - q_1| + |z_j(\theta_1 - \theta_2)| \\ &\leq |z_j||\theta_2 - \theta_1| + 4|q_2 - q_1|. \end{aligned}$$

For $q_1 < z_j < q_2$:

$$|\Phi_j(z, \theta_1, q_1) - \Phi_j(z, \theta_2, q_2)| = |\theta_1(z_j - q_1) - (1-\theta_2)(q_2 - z_j)| \leq |q_2 - q_1|.$$

For $q_2 \leq z_j$:

$$|\Phi_j(z, \theta_1, q_1) - \Phi_j(z, \theta_2, q_2)| = |\theta_1(z_j - q_1) - \theta_2(z_j - q_2)|,$$

and (5) follows along the lines of the first case.

LEMMA 4. *Expression (7) holds under the conditions of Theorem* 1.

Proof of Lemma 4: Suppose (7) was wrong. This means that there exist $\epsilon > 0, \; \delta > 0$, a subsequence M of positive integers and $(\theta^*_m)_{m \in M}$ such that

$$\text{for all } m \in M, \; \text{pr}\left\{|\Psi_m(\theta^*_m) - \Psi(\theta^*_m)| > \epsilon\right\} \geq \delta. \quad (8)$$

Since $(\theta_m)_{m \in M} \in T^M$ is bounded and at least a subsequence has a limit, there exists $\theta^* = \lim_{m\to\infty} \theta^*_m$.

Consider

$$|\Psi_m(\theta^*_m) - \Psi(\theta^*_m)| \leq |\Psi_m(\theta^*_m) - \Psi_m(\theta^*)| + |\Psi_m(\theta^*) - \Psi(\theta^*)| + |\Psi(\theta^*) - \Psi(\theta^*_m)|. \quad (9)$$

Continuity of Ψ forces the third term of the right side of (9) to converge to zero.

Regarding the second term, define a version of Ψ_n using the true quantiles instead of the empirical ones:

$$\Psi_n^*(\theta) = \frac{1}{n}\left\{\sum_{i:\ C_i=0} \mathbb{1}\left(\sum_{j=1}^{p}[\Phi_j\{Z_i,\theta,q_{1j}(\theta)\} - \Phi_j\{Z_i,\theta,q_{0j}(\theta)\}] > 0\right)\right.$$
$$\left. + \sum_{i:\ C_i=1} \mathbb{1}\left(\sum_{j=1}^{p}[\Phi_j\{Z_i,\theta,q_{1j}(\theta)\} - \Phi_j\{Z_i,\theta,q_{1j}(\theta)\}] \leq 0\right)\right\}.$$

Consider

$$|\Psi_m(\theta^*) - \Psi(\theta^*)| \leq |\Psi_m(\theta^*) - \Psi_m^*(\theta^*)| + |\Psi_m^*(\theta^*) - \Psi(\theta^*)|.$$

Because of the strong law of large numbers, $\lim_{m\to\infty} |\Psi_m^*(\theta^*) - \Psi(\theta^*)| = 0$ almost surely.

For given z and θ, Φ_j is continuous in q. Furthermore quantiles are strongly consistent, and therefore (5) will enforce $\lim_{m\to\infty} |\Psi_m(\theta^*) - \Psi_m^*(\theta^*)| = 0$ almost surely.

Now consider the first term of the right side of (9).

$$\begin{aligned}|q_{kjm}(\theta_m^*) - q_{kjm}(\theta^*)| \leq |q_{kjm}(\theta^*) - q_{kj}(\theta^*)| + |q_{kjm}(\theta_m^*) - q_{kj}(\theta_m^*)| \\ + |q_{kj}(\theta_m^*) - q_{kj}(\theta^*)|.\end{aligned} \tag{10}$$

From Theorem 3 in Mason (1982), under assumption A1, we get $\lim_{m\to\infty}\sup_{\theta\in T} |q_{kj}(\theta) - q_{kjm}(\theta)| = 0$ almost surely. This enforces the first two terms on the left side of (10) to converge to zero almost surely. The last term converges to zero because of A1. Therefore, for $m \to \infty$,

$$|q_{kjm}(\theta_m^*) - q_{kjm}(\theta^*)| \to 0 \text{ almost surely.} \tag{11}$$

Let $D_n(\theta, z) = \sum_{j=1}^{p}\{\Phi_{1jn}(z,\theta) - \Phi_{0jn}(z,\theta)\}$, $D(\theta, z) = \sum_{j=1}^{p}\{\Phi_{1j}(z,\theta) - \Phi_{0j}(z,\theta)\}$. For $\epsilon > 0$ define

$$Z_\epsilon = \{z :\ |D(\theta^*, z)| > \epsilon\} \cap \left(z :\ \sum_{j=1}^{p} |z_j| \leq \frac{1}{\epsilon}\right),$$

so that

$$\begin{aligned}|\Psi_m(\theta_m^*) - \Psi_m(\theta^*)| = \frac{1}{m}\Bigg(&\sum_{i:\ C_i=0,\ Z_i\notin Z_\epsilon}\left[\mathbb{1}_{\{D_m(\theta_m^*,Z_i)>0\}} - \mathbb{1}_{\{D_m(\theta^*,Z_i)>0\}}\right] + \\ &\sum_{i:\ C_i=1,\ Z_i\notin Z_\epsilon}\left[\mathbb{1}_{\{D_m(\theta_m^*,Z_i)\leq 0\}} - \mathbb{1}_{\{D_m(\theta^*,Z_i)\leq 0\}}\right] + \\ &\sum_{i:\ C_i=0,\ Z_i\in Z_\epsilon}\left[\mathbb{1}_{\{D_m(\theta_m^*,Z_i)>0\}} - \mathbb{1}_{\{D_m(\theta^*,Z_i)>0\}}\right] + \\ &\sum_{i:\ C_i=1,\ Z_i\in Z_\epsilon}\left[\mathbb{1}_{\{D_m(\theta_m^*,Z_i)\leq 0\}} - \mathbb{1}_{\{D_m(\theta^*,Z_i)\leq 0\}}\right]\Bigg).\end{aligned}$$

Now for large m and arbitrarily small $\nu > 0$,

$$\frac{1}{m}\left|\left(\sum_{i:\ C_i=0,\ Z_i \notin Z_\epsilon}\left[\mathbb{1}_{\{D_m(\theta^*_m, Z_i)>0\}} - \mathbb{1}_{\{D_m(\theta^*, Z_i)>0\}}\right] + \right.\right.$$
$$\left.\left.\sum_{i:\ C_i=1,\ Z_i \notin Z_\epsilon}\left[\mathbb{1}_{\{D_m(\theta^*_m, Z_i)\leq 0\}} - \mathbb{1}_{\{D_m(\theta^*, Z_i)\leq 0\}}\right]\right)\right| \leq 1 - \text{pr}(Z_\epsilon) + \nu \text{ almost surely.}$$

Furthermore, by (5),

$$|D_m(\theta^*_m, Z_i) - D_m(\theta^*, Z_i)| \leq \sum_{j=1}^{p}\{2|Z_j||\theta^*_m - \theta^*| + 8|q_{kjm}(\theta^*_m) - q_{kjm}(\theta^*)|\}\,.$$

Because $|\theta^*_m - \theta^*| \to 0$, by (11) and $\sum_{j=1}^p |Z_j| \leq 1/\epsilon$ for $Z \in Z_\epsilon$, $|D_m(\theta^*_m, Z_i) - D_m(\theta^*, Z_i)|$ becomes arbitrarily small almost surely for large enough m, and therefore for $Z_i \in Z_\epsilon$, $D_m(\theta^*_m, Z_i)$ and $D_m(\theta^*, Z_i)$ will for large enough m be either both positive or both negative, and the corresponding indicator functions will therefore be the same, almost surely.

For $\epsilon \searrow 0$, assumption A2 enforces $\text{pr}(Z_\epsilon) \to 1$. This forces the first term on the right side of (9) to zero for large m, almost surely, in contradiction to (8), which in turn proves (7).

2·2. *Theorem 2*

The assumptions for Theorem 2 are as follows.

B1. The limit $\lim_{\lambda\to\infty}\sup_{k\geq 1} E\{|U_k|\mathbb{1}_{(|U_k|>\lambda)}\} = 0$.
B2. For each $c > 0$,

$$\inf_{k\geq 1}\inf_{|x|\geq c}\inf_{\theta\in T}\left(E[\Phi_k\{U, \theta, q_k(\theta) + x\}] - E[\Phi_k\{U, \theta, q_k(\theta)\}]\right) > 0.$$

B3. For each $\epsilon > 0$,

$$\inf_{k\geq 1}\inf_{\theta\in T}\left(\min[\theta - \text{pr}\{U_k \leq q_k(\theta) - \epsilon\}, \theta - \text{pr}\{U_k \geq q_k(\theta) + \epsilon\}]\right) > 0.$$

B4. With $\mathcal{B}$ denoting the class of Borel subsets of the real line,

$$\lim_{k\to\infty}\sup_{k_1,k_2:\ |k_1-k_2|\geq k}\ \sup_{B_1,B_2\in\mathcal{B}}|\text{pr}(U_{k_1} \in B_1, U_{k_2} \in B_2) - \text{pr}(U_{k_1} \in B_1)\text{pr}(U_{k_2} \in B_2)| = 0.$$

B5. The differences $|\nu_{Xk,\theta} - \nu_{Yk,\theta}|$ are uniformly bounded.
B6. For sufficiently small $\epsilon > 0$, the proportion of values $k \in [1, p]$ for which $|\nu_{Xk,\theta} - \nu_{Yk,\theta}| > \epsilon$ for all $\theta \in T$ is bounded away from zero as p diverges.

THEOREM 2. *Assume B1–B6 hold and that both n and m diverge as $p \to \infty$. Then, with probability converging to 1 as p increases, the classifier $\mathcal{R}_{m,n,p}$ makes the correct decision, i.e.,*

$$\text{pr}\{\mathcal{R}_{m,n,p}(Z) = 1 \mid C = 0\} + \text{pr}\{\mathcal{R}_{m,n,p}(Z) = 0 \mid C = 1\} \to 0.$$

Proof of Theorem 2: In the proof of Theorem 1 in Hall et al. (2009), B2, B3, B5 and B6 enforce every statement to hold uniformly for $\theta \in T$, after definitions have been adapted to general quantile classifiers. More specifically, $W_k, D_k, D(Z), S_\lambda, d(Z), \mathcal{K}_\epsilon$ and d_k need to be defined as functions of θ with quantiles replacing medians, Φ_k replacing the absolute value where B2 is applied and $q_k(\theta)$ replacing zero where B3 is applied. Equations (A.1)–(A.6) in Hall et al. (2009) then hold uniformly over T.

Remark 1. We expect that similar arguments can be used to prove a version of Theorem 2 in Hall et al. (2009), which has different assumptions, for the quantile classifier.

3. Detailed results of the simulation study

3·1. *Software information and tuning parameters*

The package `quantileDA` available on CRAN/R computes quantile, median and centroid classifiers.

The implementation of the classification methods used for comparison has been done with the following software and parameter choices. We used the R package `MASS` to implement Fisher's LDA and the package `pamr` for the nearest-shrunken centroid. The k-nearest-neighbor classifier has been run under the library `class` with a number of neighbors, k, chosen between 1 and 9 by cross-validation in the training set. For the support vector machine we used the library `e1071` with the combination of tuning parameters $\gamma = (0{\cdot}001, 0{\cdot}01, 0{\cdot}1, 1, 2)$ and `Cost`$= (1, 2, 4, 8, 16)$ optimized in each dataset according to the function `tune.svm`. For the classification trees we used the package `rpart` with the combination of tuning parameters `minsplit`$= (5, 10, 15, 20)$ and `cp`$= (0{\cdot}001, 0{\cdot}01, 0{\cdot}1, 0{\cdot}2)$ selected by cross-validation. The naive Bayes classifier has been fitted by the package `klaR` with kernel estimate densities instead of Gaussian densities in order to gain flexibility. We used the package `stepPlr` for penalized logistic regression with regularization parameter, λ, chosen by cross-validation in $\lambda = (0{\cdot}00001, 0{\cdot}0001, 0{\cdot}001, 0{\cdot}01, 0{\cdot}1, 1)$.

3·2. *Classification results*

The tables show the average misclassification rates multiplied by 100 and the average of the optimal θ values across all the 100 data sets in each setting considered. Standard errors are reported in brackets. Tables 3 and 4 also report the misclassification rates obtained by the median classifier on data preprocessed by the Box–Cox power transformation in order to check whether this transformation could improve the classification performance of the median classifier by reducing the skewness of variables. For each variable, we considered 100 values equally spaced in the interval $[-2, 2]$ for the Box–Cox parameter, and the single parameter value that maximized the profile log-likelihood of a Gaussian distribution for each variable was selected in the training set and applied to the test set. The Box–Cox transformation does not improve the classification performance, because the most discriminative information between the two classes in the tails of the class distributions still remains located in the tails even if the distributional shape changes.

References

Hall, P., Titterington, D. M. & Xue, J. H. (2009). Median-based classifiers for high-dimensional data. *Journal of the American Statistical Society* **104**, 1597–1608.

Mason, D. M. (1982). Some Characterizations of Almost Sure Bounds for Weighted Multidimensional Empirical Distributions and a Glivenko–Cantelli Theorem for Sample Quantiles. *Zeitschrift fuer Wahrscheinlichkeitstheorie und verwandte Gebiete* **50**, 505–513.

Table 1. *Simulation study: independent identically distributed symmetric variables. Misclassification rates multiplied by 100 (with standard errors in brackets; all rounded to one digit after the decimal point) for different methods. Rows 2 and 4 contain the mean of the chosen values of θ in the training sets.*

	$p=50$			$p=100$			$p=500$		
$n=50$	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	17·0 (0·6)	27·5 (0·6)	42·2 (0·6)	10·1 (0·5)	20·0 (0·7)	41·2 (0·7)	1·5 (0·2)	5·6 (0·4)	31·5 (0·8)
$\bar{\theta}$ Galton	46·1 (1·8)	44·4 (1·8)	44·3 (1·7)	45·7 (1·3)	44·0 (1·4)	42·7 (1·1)	48·3 (0·3)	48·3 (0·2)	48·3 (0·3)
QCS	16·6 (0·8)	28·3 (0·7)	42·5 (0·6)	9·6 (0·6)	20·7 (0·7)	41·1 (0·6)	1·6 (0·2)	5·8 (0·3)	31·4 (0·7)
$\bar{\theta}$ Skewn.	49·4 (1·0)	41·3 (1·8)	42·9 (1·9)	39·5 (1·2)	42·7 (1·5)	43·6 (1·5)	42·8 (0·6)	43·1 (0·3)	44·1 (0·3)
CC	16·2 (0·8)	27·3 (0·7)	43·1 (0·5)	10·5 (0·7)	22·0 (1·0)	41·9 (0·6)	3·5 (0·8)	13·3 (1·4)	37·1 (0·8)
MC	16·5 (0·5)	26·9 (0·6)	42·1 (0·5)	9·7 (0·5)	19·3 (0·6)	41·9 (0·5)	1·7 (0·2)	6·4 (0·4)	32·0 (0·7)
LDA	37·5 (0·7)	41·2 (0·7)	43·5 (0·5)	22·8 (0·7)	30·4 (0·7)	42·7 (0·5)	26·0 (0·9)	35·8 (0·8)	43·2 (0·5)
knn	19·1 (0·7)	31·1 (0·8)	44·2 (0·4)	14·1 (0·7)	29·0 (0·8)	43·4 (0·6)	5·6 (0·7)	21·1 (1·3)	42·8 (0·6)
n-Bayes	26·7 (0·7)	35·5 (0·8)	43·3 (0·5)	20·2 (0·7)	33·2 (0·8)	43·5 (0·5)	14·1 (0·9)	27·4 (1·0)	43·4 (0·5)
SVM	18·0 (0·9)	25·6 (0·8)	42·6 (0·5)	9·9 (0·5)	17·5 (0·5)	43·1 (0·4)	3·9 (0·3)	5·9 (0·5)	26·7 (0·7)
NCS	29·4 (0·8)	36·3 (0·8)	42·0 (0·6)	23·0 (0·7)	31·2 (0·8)	40·7 (0·6)	10·4 (0·6)	18·4 (0·8)	36·3 (0·8)
stepPlr	14·7 (0·5)	23·5 (0·7)	41·8 (0·6)	6·9 (0·4)	15·3 (0·5)	38·3 (0·6)	1·1 (0·1)	3·1 (0·3)	25·1 (0·7)
rpart	41·3 (0·6)	42·4 (0·6)	44·3 (0·4)	38·8 (0·6)	42·1 (0·5)	41·7 (0·6)	39·2 (0·6)	39·2 (0·6)	42·1 (0·6)
$n=100$	$p=50$			$p=100$			$p=500$		
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	15·4 (0·4)	25·2 (0·5)	41·9 (0·4)	9·3 (0·4)	17·5 (0·4)	39·6 (0·5)	1·4 (0·1)	3·8 (0·2)	26·1 (0·4)
$\bar{\theta}$ Galton	43·3 (1·5)	43·2 (1·6)	42·2 (1·5)	47·0 (1·8)	44·2 (1·6)	43·8 (1·3)	49·1 (0·5)	48·2 (0·4)	47·9 (0·5)
QCS	15·5 (0·4)	25·1 (0·4)	42·5 (0·4)	9·7 (0·4)	17·8 (0·5)	39·7 (0·5)	1·5 (0·1)	3·9 (0·2)	25·5 (0·4)
$\bar{\theta}$ Skewn.	43·5 (1·7)	47·2 (1·6)	46·9 (1·9)	41·9 (1·7)	45·1 (1·6)	47·1 (1·4)	45·5 (0·5)	46·1 (0·3)	47·1 (0·5)
CC	13·0 (0·6)	21·8 (0·6)	42·4 (0·5)	6·7 (0·3)	16·1 (0·6)	37·4 (0·6)	1·0 (0·1)	4·4 (0·5)	30·1 (0·8)
MC	13·8 (0·4)	23·0 (0·5)	41·5 (0·5)	7·6 (0·3)	16·5 (0·3)	36·6 (0·5)	1·3 (0·1)	3·9 (0·2)	26·4 (0·5)
LDA	18·0 (0·5)	27·2 (0·5)	42·6 (0·4)	34·6 (0·8)	38·6 (0·6)	44·7 (0·4)	12·4 (0·4)	21·6 (0·5)	41·9 (0·5)
knn	15·3 (0·4)	26·3 (0·5)	43·6 (0·5)	9·1 (0·4)	21·1 (0·7)	43·1 (0·5)	2·1 (0·2)	13·2 (1·0)	40·6 (0·6)
n-Bayes	21·2 (0·5)	31·7 (0·6)	44·3 (0·4)	16·3 (0·5)	27·8 (0·7)	44·3 (0·4)	9·6 (0·5)	19·8 (0·9)	41·9 (0·6)
SVM	11·8 (0·3)	20·7 (0·4)	41·9 (0·5)	6·9 (0·3)	13·8 (0·3)	35·8 (0·6)	2·6 (0·2)	3·6 (0·2)	21·1 (0·5)
NCS	24·8 (0·6)	29·9 (0·5)	41·1 (0·5)	17·4 (0·4)	25·9 (0·6)	38·2 (0·6)	4·4 (0·3)	9·8 (0·4)	28·3 (0·7)
stepPlr	12·9 (0·3)	22·5 (0·5)	40·7 (0·4)	6·6 (0·3)	13·9 (0·3)	36·7 (0·5)	0·8 (0·1)	2·3 (0·2)	19·6 (0·4)
rpart	41·7 (0·4)	42·1 (0·4)	43·8 (0·5)	41·3 (0·5)	41·7 (0·5)	43·2 (0·5)	41·5 (0·4)	40·0 (0·4)	41·9 (0·4)
$n=500$	$p=50$			$p=100$			$p=500$		
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	12·7 (0·2)	20·0 (0·2)	37·6 (0·2)	7·1 (0·1)	13·1 (0·1)	32·5 (0·2)	1·2 (0·1)	3·1 (0·1)	17·0 (0·2)
$\bar{\theta}$ Galton	42·7 (1·1)	42·6 (1·0)	41·9 (1·0)	44·3 (1·3)	44·3 (1·1)	43·2 (1·0)	48·2 (2·0)	45·2 (1·9)	44·3 (1·0)
QCS	12·7 (0·2)	19·8 (0·2)	37·5 (0·3)	7·1 (0·1)	13·0 (0·1)	32·5 (0·2)	1·2 (0·1)	3·0 (0·1)	17·4 (0·2)
$\bar{\theta}$ Skewn.	52·6 (1·2)	50·3 (1·0)	48·5 (1·3)	48·4 (1·3)	49·8 (1·2)	50·1 (1·1)	49·3 (1·7)	51·4 (1·7)	49·0 (1·3)
CC	9·4 (0·1)	16·7 (0·2)	35·7 (0·3)	5·6 (0·4)	10·0 (0·1)	30·4 (0·2)	0·7 (0·0)	1·8 (0·1)	15·9 (0·6)
MC	12·2 (0·2)	19·4 (0·2)	36·3 (0·2)	6·9 (0·1)	12·5 (0·2)	31·7 (0·2)	1·1 (0·0)	2·5 (0·1)	16·2 (0·2)
LDA	11·2 (0·2)	18·6 (0·2)	35·7 (0·3)	7·1 (0·1)	12·8 (0·2)	31·6 (0·2)	33·8 (0·6)	37·9 (0·5)	45·5 (0·3)
knn	10·8 (0·1)	20·0 (0·2)	41·4 (0·2)	5·8 (0·1)	13·2 (0·2)	39·3 (0·3)	0·9 (0·0)	4·0 (0·2)	31·4 (0·6)
n-Bayes	14·2 (0·2)	23·3 (0·3)	41·9 (0·3)	9·8 (0·1)	18·3 (0·3)	39·6 (0·4)	5·3 (0·2)	9·7 (0·4)	35·1 (0·7)
SVM	10·8 (0·1)	16·5 (0·2)	34·1 (0·2)	4·9 (0·1)	11·5 (0·1)	30·1 (0·2)	1·4 (0·1)	2·5 (0·1)	12·9 (0·2)
NCS	11·7 (0·2)	18·8 (0·2)	33·3 (0·3)	7·0 (0·3)	11·7 (0·2)	28·0 (0·2)	1·0 (0·0)	2·4 (0·1)	12·7 (0·2)
stepPlr	12·0 (0·2)	19·0 (0·2)	36·0 (0·2)	6·6 (0·1)	13·8 (0·2)	31·4 (0·2)	0·7 (0·0)	1·8 (0·1)	14·3 (0·2)
rpart	41·2 (0·2)	41·2 (0·3)	41·4 (0·2)	40·9 (0·2)	41·5 (0·2)	41·1 (0·2)	41·5 (0·3)	40·9 (0·2)	41·2 (0·2)

Table 2. *Simulation study: dependent identically distributed symmetric variables. Misclassification rates multiplied by 100 (with standard errors in brackets; all rounded to one digit after the decimal point) for different methods. Rows 2 and 4 contain the mean of the chosen values of θ in the training sets.*

				$n = 50$					
	$p = 50$			$p = 100$			$p = 500$		
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	19·9 (0·7)	30·0 (0·8)	42·6 (0·6)	11·2 (0·6)	21·9 (0·7)	41·6 (0·6)	2·2 (0·2)	7·7 (0·4)	34·4 (0·7)
$\bar{\theta}$ Galton	44·7 (1·9)	43·5 (1·9)	45·9 (2·3)	46·7 (1·4)	44·3 (1·6)	41·8 (1·5)	47·0 (0·6)	48·0 (0·3)	48·1 (0·3)
QCS	19·3 (0·6)	30·1 (0·8)	42·6 (0·6)	11·1 (0·5)	23·6 (0·7)	41·2 (0·6)	3·3 (0·3)	9·6 (0·5)	35·0 (0·6)
$\bar{\theta}$ Skewn.	51·5 (1·0)	50·1 (1·6)	49·7 (2·4)	50·9 (0·5)	52·2 (1·4)	51·5 (1·8)	50·1 (0·4)	50·8 (0·6)	49·3 (0·3)
CC	19·0 (0·7)	31·4 (0·8)	42·5 (0·6)	13·5 (0·9)	24·6 (0·9)	42·2 (0·6)	3·3 (0·6)	12·3 (1·0)	39·0 (0·7)
MC	17·9 (0·6)	28·6 (0·7)	42·0 (0·6)	10·3 (0·4)	21·5 (0·6)	40·9 (0·7)	2·0 (0·2)	7·8 (0·4)	34·8 (0·6)
LDA	16·0 (1·0)	21·2 (0·9)	36·2 (0·9)	23·6 (0·7)	33·2 (0·8)	42·9 (0·5)	31·1 (0·8)	36·9 (0·7)	43·6 (0·5)
knn	21·4 (0·7)	33·1 (0·7)	43·6 (0·5)	13·8 (0·6)	28·5 (0·8)	44·0 (0·5)	8·2 (0·9)	20·9 (1·2)	44·6 (0·5)
n-Bayes	26·3 (0·8)	36·4 (0·7)	44·5 (0·4)	21·0 (0·7)	33·7 (0·8)	43·4 (0·5)	14·8 (0·9)	27·7 (1·0)	43·6 (0·5)
SVM	9·3 (0·4)	18·3 (0·6)	41·0 (0·6)	7·7 (0·5)	14·8 (0·6)	38·3 (0·7)	3·6 (0·3)	6·5 (0·4)	32·8 (0·8)
NCS	30·7 (0·7)	36·1 (0·7)	43·2 (0·5)	26·4 (0·7)	31·5 (0·8)	41·4 (0·6)	11·3 (0·6)	18·3 (0·7)	38·1 (0·7)
stepPlr	8·7 (0·4)	16·8 (0·6)	37·5 (0·7)	5·3 (0·3)	11·7 (0·5)	36·7 (0·8)	1·2 (0·1)	4·6 (0·3)	29·7 (0·6)
rpart	38·6 (0·6)	40·6 (0·6)	43·1 (0·5)	40·1 (0·6)	39·9 (0·6)	43·2 (0·5)	39·0 (0·7)	40·3 (0·6)	42·1 (0·6)
				$n = 100$					
	$p = 50$			$p = 100$			$p = 500$		
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	16·1 (0·4)	25·2 (0·5)	43·4 (0·4)	9·4 (0·4)	17·6 (0·5)	39·6 (0·6)	1·5 (0·1)	4·6 (0·2)	29·0 (0·4)
$\bar{\theta}$ Galton	41·0 (1·7)	42·8 (1·4)	45·1 (2·0)	46·1 (1·7)	43·9 (1·5)	41·1 (1·4)	48·7 (0·3)	48·5 (0·7)	46·6 (0·9)
QCS	15·3 (0·4)	25·3 (0·6)	43·8 (0·4)	8·8 (0·3)	17·6 (0·4)	39·5 (0·6)	1·8 (0·1)	5·6 (0·4)	29·8 (0·5)
$\bar{\theta}$ Skewn.	50·0 (0·6)	51·7 (1·0)	49·1 (2·1)	49·9 (0·4)	50·7 (0·8)	48·7 (1·6)	50·0 (0·2)	50·4 (0·6)	48·6 (1·1)
CC	15·5 (0·6)	25·1 (0·5)	44·0 (0·4)	8·8 (0·5)	18·6 (0·7)	41·2 (0·5)	1·4 (0·1)	6·9 (0·6)	33·4 (0·7)
MC	14·8 (0·4)	24·4 (0·5)	43·1 (0·5)	8·3 (0·3)	16·4 (0·4)	39·4 (0·6)	1·5 (0·1)	4·6 (0·2)	28·6 (0·4)
LDA	0·6 (0·1)	1·9 (0·1)	4·2 (0·2)	9·0 (0·8)	11·9 (0·9)	29·0 (0·9)	15·9 (0·4)	25·8 (0·6)	41·9 (0·5)
knn	15·2 (0·4)	27·4 (0·5)	44·4 (0·4)	12·3 (0·4)	26·0 (0·5)	43·6 (0·4)	3·3 (0·3)	13·8 (0·8)	41·1 (0·5)
n-Bayes	21·4 (0·6)	32·5 (0·6)	44·9 (0·4)	17·2 (0·6)	28·1 (0·8)	44·1 (0·4)	10·4 (0·5)	22·6 (1·0)	42·9 (0·6)
SVM	5·9 (0·3)	11·4 (0·4)	42·3 (0·4)	4·0 (0·2)	6·7 (0·3)	32·8 (0·6)	2·3 (0·2)	4·4 (0·2)	23·4 (0·5)
NCS	27·6 (0·6)	32·1 (0·5)	42·7 (0·6)	20·7 (0·4)	27·2 (0·6)	38·9 (0·6)	6·4 (0·3)	11·6 (0·4)	29·2 (0·7)
stepPlr	4·4 (0·2)	9·9 (0·3)	21·4 (0·6)	3·7 (0·2)	6·6 (0·3)	29·1 (0·5)	0·7 (0·1)	2·6 (0·2)	21·3 (0·4)
rpart	41·7 (0·5)	41·6 (0·4)	42·8 (0·5)	40·9 (0·4)	41·8 (0·5)	41·7 (0·5)	39·0 (0·4)	41·7 (0·5)	41·9 (0·5)
				$n = 500$					
	$p = 50$			$p = 100$			$p = 500$		
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	13·0 (0·2)	21·0 (0·2)	38·3 (0·3)	6·8 (0·1)	12·1 (0·1)	33·0 (0·2)	1·2 (0·1)	3·2 (0·1)	19·0 (0·2)
$\bar{\theta}$ Galton	42·0 (1·1)	42·4 (1·1)	45·7 (1·3)	42·6 (1·2)	42·0 (0·9)	44·2 (1·1)	47·2 (1·7)	45·5 (1·7)	42·6 (1·1)
QCS	12·9 (0·2)	20·7 (0·2)	38·4 (0·3)	6·6 (0·1)	11·9 (0·2)	33·1 (0·3)	1·2 (0·1)	3·1 (0·1)	18·5 (0·2)
$\bar{\theta}$ Skewn.	50·2 (0·4)	50·1 (0·5)	49·7 (1·4)	50·0 (0·2)	49·6 (0·3)	50·8 (1·0)	50·0 (0·1)	50·0 (0·1)	49·5 (0·7)
CC	10·2 (0·2)	18·4 (0·3)	38·3 (0·3)	4·7 (0·1)	9·6 (0·2)	33·3 (0·3)	0·7 (0·0)	2·3 (0·1)	19·1 (0·3)
MC	12·6 (0·1)	20·5 (0·2)	37·8 (0·2)	6·3 (0·1)	11·7 (0·2)	32·4 (0·2)	1·1 (0·0)	2·9 (0·1)	18·1 (0·2)
LDA	0·2 (0·0)	0·7 (0·0)	1·9 (0·1)	0·0 (0·0)	0·0 (0·0)	0·5 (0·0)	1·2 (0·1)	6·0 (0·5)	19·5 (0·7)
knn	11·4 (0·1)	20·8 (0·2)	41·2 (0·2)	5·6 (0·1)	12·8 (0·3)	39·0 (0·2)	1·1 (0·1)	4·9 (0·2)	32·5 (0·6)
n-Bayes	14·8 (0·2)	23·5 (0·3)	42·8 (0·3)	9·8 (0·2)	17·4 (0·3)	40·2 (0·4)	5·6 (0·2)	10·2 (0·4)	35·9 (0·8)
SVM	1·8 (0·1)	4·2 (0·1)	15·4 (0·2)	1·2 (0·0)	2·0 (0·1)	14·8 (0·2)	1·3 (0·1)	2·0 (0·1)	8·7 (0·1)
NCS	14·3 (0·3)	20·1 (0·3)	34·3 (0·3)	7·5 (0·2)	11·8 (0·2)	27·8 (0·3)	1·3 (0·1)	2·9 (0·1)	12·8 (0·2)
stepPlr	1·8 (0·1)	2·5 (0·1)	7·6 (0·1)	1·2 (0·1)	1·9 (0·1)	8·7 (0·1)	0·4 (0·0)	1·3 (0·1)	8·4 (0·1)
rpart	41·5 (0·2)	41·2 (0·2)	41·9 (0·3)	41·7 (0·2)	41·5 (0·2)	41·4 (0·2)	40·6 (0·2)	41·6 (0·2)	41·0 (0·2)

Table 3. *Simulation study: independent identically distributed asymmetric variables. Misclassification rates multiplied by 100 (with standard errors in brackets; all rounded to one digit after the decimal point) for different methods. Rows 2 and 4 contain the mean of the chosen values of θ in the training sets.*

					$n = 50$				
		$p = 50$			$p = 100$			$p = 500$	
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	25·4 (0·9)	35·7 (0·8)	43·3 (0·5)	25·7 (1·0)	35·6 (0·9)	44·3 (0·5)	25·8 (0·7)	35·5 (0·7)	44·6 (0·4)
$\bar{\theta}$ Galton	18·2 (1·6)	27·5 (2·6)	46·2 (3·1)	34·9 (2·7)	44·3 (2·8)	60·4 (2·6)	48·2 (2·3)	52·5 (2·0)	61·1 (1·1)
QCS	19·6 (0·7)	27·7 (0·8)	42·0 (0·5)	21·3 (0·8)	23·8 (0·7)	41·5 (0·6)	26·7 (1·0)	25·6 (0·7)	30·3 (0·9)
$\bar{\theta}$ Skewn.	5·9 (0·5)	7·7 (1·0)	29·1 (3·0)	6·4 (1·0)	9·7 (1·8)	38·1 (3·8)	5·5 (1·0)	4·7 (0·9)	14·8 (3·0)
CC	43·2 (0·5)	44·2 (0·5)	45·5 (0·4)	42·7 (0·5)	43·1 (0·5)	44·0 (0·5)	38·6 (0·6)	42·6 (0·5)	45·1 (0·4)
MC	37·7 (0·7)	43·2 (0·5)	44·0 (0·5)	34·3 (0·7)	40·2 (0·6)	44·8 (0·4)	16·8 (0·6)	29·8 (0·7)	43·0 (0·5)
(Box–Cox) MC	37·0 (0·6)	42·2 (0·6)	44·7 (0·4)	33·6 (0·7)	40·4 (0·7)	44·3 (0·4)	18·9 (0·6)	30·0 (0·7)	43·7 (0·5)
LDA	43·6 (0·5)	44·5 (0·4)	44·2 (0·5)	44·0 (0·4)	43·4 (0·4)	44·9 (0·4)	43·1 (0·5)	43·8 (0·5)	45·1 (0·4)
knn	45·0 (0·4)	44·1 (0·5)	45·0 (0·4)	45·6 (0·4)	44·0 (0·4)	45·3 (0·4)	44·7 (0·4)	45·3 (0·4)	44·8 (0·4)
n-Bayes	42·5 (0·5)	44·0 (0·5)	44·4 (0·4)	42·7 (0·5)	43·7 (0·5)	45·1 (0·4)	39·8 (0·7)	43·3 (0·5)	45·2 (0·4)
SVM	43·0 (0·5)	43·5 (0·5)	44·4 (0·4)	42·4 (0·5)	44·1 (0·4)	44·2 (0·4)	35·6 (0·7)	41·4 (0·6)	45·2 (0·5)
NCS	44·6 (0·4)	44·9 (0·4)	44·9 (0·4)	44·8 (0·5)	44·2 (0·5)	44·3 (0·4)	42·8 (0·6)	42·6 (0·5)	45·3 (0·4)
stepPlr	43·4 (0·5)	43·8 (0·5)	44·5 (0·4)	42·3 (0·5)	43·7 (0·5)	44·4 (0·4)	36·0 (0·7)	41·5 (0·6)	44·2 (0·4)
rpart	41·7 (0·6)	42·3 (0·5)	43·9 (0·5)	42·6 (0·5)	43·6 (0·5)	44·3 (0·4)	41·5 (0·7)	42·7 (0·5)	43·8 (0·5)
					$n = 100$				
		$p = 50$			$p = 100$			$p = 500$	
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	9·4 (0·4)	18·4 (0·5)	42·0 (0·5)	7·0 (0·4)	14·1 (0·5)	37·5 (0·6)	5·0 (0·7)	14·4 (1·1)	42·4 (0·6)
$\bar{\theta}$ Galton	3·5 (0·2)	3·6 (0·3)	18·7 (2·5)	5·4 (0·6)	5·9 (0·6)	16·6 (2·3)	26·8 (2·0)	28·7 (2·4)	56·4 (2·5)
QCS	8·7 (0·3)	16·9 (0·4)	40·9 (0·5)	6·2 (0·3)	11·8 (0·4)	35·3 (0·5)	1·1 (0·2)	6·4 (0·7)	26·6 (1·0)
$\bar{\theta}$ Skewn.	3·6 (0·2)	3·1 (0·2)	13·5 (2·1)	2·8 (0·2)	3·9 (0·2)	9·8 (1·7)	17·1 (1·5)	10·7 (1·7)	17·0 (3·1)
CC	42·5 (0·5)	44·5 (0·3)	46·0 (0·3)	41·3 (0·5)	44·3 (0·4)	46·5 (0·3)	33·2 (0·5)	39·8 (0·5)	45·9 (0·3)
MC	34·5 (0·6)	41·0 (0·5)	45·7 (0·4)	29·8 (0·4)	37·9 (0·6)	45·4 (0·3)	11·4 (0·3)	24·9 (0·4)	43·3 (0·4)
(Box–Cox) MC	34·4 (0·5)	41·5 (0·5)	45·9 (0·3)	28·5 (0·5)	37·0 (0·6)	44·4 (0·3)	13·3 (0·3)	26·4 (0·5)	43·9 (0·4)
LDA	43·9 (0·4)	45·7 (0·3)	46·1 (0·3)	45·8 (0·3)	45·4 (0·4)	46·2 (0·3)	41·4 (0·5)	44·2 (0·4)	45·5 (0·3)
knn	45·8 (0·4)	45·7 (0·3)	45·9 (0·3)	45·8 (0·3)	46·1 (0·3)	45·6 (0·3)	44·0 (0·4)	46·4 (0·3)	46·5 (0·3)
n-Bayes	41·9 (0·5)	44·7 (0·4)	46·0 (0·3)	41·4 (0·5)	44·3 (0·4)	45·5 (0·3)	37·3 (0·4)	42·2 (0·4)	45·9 (0·3)
SVM	45·6 (0·3)	44·6 (0·4)	46·6 (0·3)	41·6 (0·4)	44·1 (0·4)	45·9 (0·3)	30·6 (0·5)	39·4 (0·5)	45·7 (0·3)
NCS	45·5 (0·3)	46·0 (0·3)	46·5 (0·3)	44·5 (0·4)	45·6 (0·4)	46·4 (0·3)	41·6 (0·5)	44·0 (0·4)	45·6 (0·3)
stepPlr	44·2 (0·4)	45·8 (0·3)	46·2 (0·3)	43·0 (0·4)	45·0 (0·4)	45·6 (0·3)	32·6 (0·5)	40·4 (0·5)	45·5 (0·3)
rpart	37·7 (0·6)	43·5 (0·4)	45·0 (0·4)	42·3 (0·4)	42·7 (0·5)	44·4 (0·4)	41·3 (0·5)	41·9 (0·5)	42·7 (0·5)
					$n = 500$				
		$p = 50$			$p = 100$			$p = 500$	
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	2·1 (0·1)	8·6 (0·1)	33·9 (0·3)	0·2 (0·0)	2·6 (0·1)	25·7 (0·2)	0·0 (0·0)	0·0 (0·0)	6·3 (0·1)
$\bar{\theta}$ Galton	2·1 (0·1)	2·4 (0·1)	4·8 (0·3)	2·2 (0·1)	2·2 (0·1)	3·4 (0·2)	17·5 (0·4)	2·0 (0·0)	2·7 (0·1)
QCS	2·1 (0·1)	8·6 (0·1)	33·9 (0·3)	0·2 (0·0)	2·6 (0·1)	25·7 (0·2)	0·0 (0·0)	0·0 (0·0)	6·3 (0·1)
$\bar{\theta}$ Skewn.	2·1 (0·1)	2·4 (0·1)	4·8 (0·3)	2·2 (0·1)	2·2 (0·1)	3·4 (0·2)	17·5 (0·4)	2·0 (0·0)	2·7 (0·1)
CC	40·2 (0·2)	43·7 (0·2)	47·8 (0·2)	36·4 (0·2)	41·5 (0·2)	47·6 (0·2)	22·4 (0·2)	32·7 (0·2)	46·0 (0·2)
MC	30·7 (0·2)	37·1 (0·2)	45·8 (0·2)	23·2 (0·2)	32·1 (0·2)	45·2 (0·2)	5·2 (0·1)	15·1 (0·2)	38·5 (0·2)
(Box–Cox) MC	29·9 (0·2)	37·1 (0·2)	46·0 (0·2)	23·5 (0·2)	32·6 (0·2)	45·0 (0·3)	5·4 (0·1)	15·7 (0·2)	39·7 (0·2)
LDA	40·7 (0·2)	43·9 (0·2)	47·8 (0·2)	37·2 (0·2)	42·1 (0·2)	47·9 (0·2)	46·8 (0·2)	47·5 (0·2)	48·3 (0·2)
knn	45·5 (0·2)	47·5 (0·2)	48·1 (0·1)	45·2 (0·2)	47·2 (0·2)	48·3 (0·1)	45·6 (0·2)	47·3 (0·2)	48·5 (0·1)
n-Bayes	32·3 (0·2)	38·8 (0·2)	47·0 (0·2)	32·0 (0·2)	38·5 (0·2)	46·9 (0·2)	28·3 (0·2)	36·7 (0·2)	46·7 (0·2)
SVM	35·5 (0· 3)	41·6 (0·2)	47·7 (0·2)	31·9 (0·3)	39·7 (0·2)	47·5 (0·2)	19·2 (0·2)	31·9 (0·2)	46·0 (0·2)
NCS	44·9 (0·2)	46·4 (0·2)	48·0 (0·2)	43·1 (0·2)	45·4 (0·2)	47·7 (0·2)	35·8 (0·3)	40·2 (0·2)	46·3 (0·2)
stepPlr	40·3 (0·2)	43·7 (0·2)	47·8 (0·2)	37·2 (0·2)	42·3 (0·2)	47·6 (0·2)	27·8 (0·2)	36·2 (0·2)	46·8 (0·2)
rpart	20·7 (0·3)	20·3 (0·3)	35·5 (0·3)	20·5 (0·3)	21·5 (0·3)	27·5 (0·4)	23·1 (0·3)	22·0 (0·3)	21·7 (0·3)

Table 4. *Simulation study: dependent identically distributed asymmetric variables. Misclassification rates multiplied by 100 (with standard errors in brackets; all rounded to one digit after the decimal point) for different methods. Rows 2 and 4 contain the mean of the chosen values of θ in the training sets.*

					$n = 50$				
		$p = 50$			$p = 100$			$p = 500$	
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	25·6 (1·0)	34·7 (0·8)	44·2 (0·5)	24·7 (1·0)	34·1 (0·9)	43·9 (0·5)	27·2 (0·9)	37·3 (0·8)	44·0 (0·5)
$\bar{\theta}$ Galton	16·0 (2·2)	26·7 (3·1)	46·3 (3·4)	24·9 (2·2)	32·4 (3·1)	65·2 (3·3)	45·7 (2·7)	53·4 (2·4)	65·2 (1·6)
QCS	21·3 (0·7)	27·1 (0·7)	43·2 (0·5)	19·9 (0·9)	25·6 (0·7)	40·9 (0·7)	27·0 (1·0)	26·3 (0·9)	33·4 (0·8)
$\bar{\theta}$ Skewn.	7·7 (1·4)	5·2 (0·7)	28·7 (3·7)	5·0 (0·5)	7·6 (1·7)	31·4 (3·8)	7·2 (1·6)	7·8 (2·0)	21·7 (3·7)
CC	44·1 (0·5)	44·8 (0·4)	44·3 (0·4)	42·7 (0·5)	43·7 (0·5)	44·4 (0·4)	37·6 (0·7)	42·8 (0·5)	44·5 (0·4)
MC	39·3 (0·6)	42·4 (0·5)	44·9 (0·4)	35·1 (0·7)	42·1 (0·5)	44·7 (0·4)	20·7 (0·6)	33·0 (0·6)	44·0 (0·5)
(Box–Cox) MC	41·4 (0·6)	39·1 (0·7)	42·9 (0·5)	34·8 (0·6)	41·2 (0·6)	44·0 (0·5)	20·9 (0·6)	32·3 (0·7)	43·3 (0·5)
LDA	44·3 (0·5)	43·9 (0·4)	44·6 (0·4)	43·9 (0·5)	44·0 (0·5)	44·7 (0·5)	43·0 (0·5)	44·5 (0·4)	44·2 (0·4)
knn	43·7 (0·5)	44·7 (0·4)	45·2 (0·4)	43·8 (0·4)	43·9 (0·5)	44·7 (0·4)	44·4 (0·5)	44·2 (0·5)	46·0 (0·4)
n-Bayes	43·0 (0·6)	44·0 (0·5)	44·5 (0·4)	43·0 (0·5)	44·0 (0·5)	43·9 (0·4)	39·7 (0·6)	42·8 (0·5)	44·6 (0·4)
SVM	42·1 (0·6)	44·1 (0·4)	45·4 (0·4)	42·8 (0·6)	44·6 (0·4)	44·0 (0·5)	37·6 (0·6)	42·6 (0·5)	44·7 (0·4)
NCS	44·3 (0·4)	45·0 (0·4)	45·3 (0·4)	44·1 (0·4)	43·7 (0·5)	44·6 (0·4)	41·7 (0·6)	43·1 (0·5)	44·8 (0·4)
stepPlr	43·5 (0·5)	43·9 (0·5)	44·1 (0·4)	42·5 (0·5)	43·2 (0·5)	44·6 (0·4)	36·8 (0·7)	42·6 (0·5)	44·1 (0·5)
rpart	40·5 (0·7)	42·2 (0·5)	44·9 (0·4)	38·8 (0·7)	41·5 (0·7)	43·9 (0·5)	41·3 (0·6)	42·4 (0·5)	43·9 (0·4)
					$n = 100$				
		$p = 50$			$p = 100$			$p = 500$	
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	12·2 (0·5)	19·6 (0·6)	41·3 (0·5)	8·2 (0·4)	14·8 (0·5)	39·5 (0·6)	2·2 (0·3)	9·2 (0·8)	37·6 (1·0)
$\bar{\theta}$ Galton	4·3 (0·4)	4·0 (0·4)	18·8 (3·0)	4·2 (0·4)	4·4 (0·4)	25·6 (3·6)	13·4 (1·3)	12·0 (1·7)	50·7 (3·9)
QCS	10·6 (0·4)	18·1 (0·5)	40·4 (0·5)	6·8 (0·3)	13·0 (0·5)	36·2 (0·5)	1·1 (0·1)	6·6 (0·3)	24·8 (0·7)
$\bar{\theta}$ Skewn.	3·5 (0·3)	3·0 (0·2)	13·8 (2·7)	3·1 (0·2)	3·6 (0·2)	10·0 (2·2)	7·6 (1·0)	4·0 (0·6)	8·9 (2·4)
CC	43·4 (0·4)	45·2 (0·4)	46·0 (0·3)	43·0 (0·5)	45·1 (0·3)	45·9 (0·3)	34·7 (0·5)	41·8 (0·5)	45·8 (0·3)
MC	36·0 (0·5)	42·3 (0·5)	45·5 (0·3)	31·5 (0·5)	41·0 (0·5)	46·0 (0·3)	13·6 (0·3)	28·4 (0·4)	43·8 (0·4)
(Box–Cox) MC	44·0 (0·5)	36·2 (0·5)	42·4 (0·5)	30·5 (0·6)	39·1 (0·5)	44·9 (0·4)	15·3 (0·4)	29·0 (0·5)	44·5 (0·4)
LDA	43·5 (0·5)	44·4 (0·4)	45·5 (0·3)	44·9 (0·3)	45·8 (0·3)	46·5 (0·3)	41·9 (0·5)	44·7 (0·4)	45·8 (0·3)
knn	46·4 (0·3)	46·2 (0·3)	46·4 (0·3)	44·7 (0·3)	46·0 (0·3)	46·4 (0·3)	44·9 (0·4)	46·6 (0·3)	46·4 (0·3)
n-Bayes	43·0 (0·4)	44·3 (0·4)	45·8 (0·3)	42·2 (0·4)	45·2 (0·4)	46·2 (0·3)	37·1 (0·4)	43·6 (0·4)	46·1 (0·3)
SVM	39·2 (0·5)	44·4 (0·4)	45·6 (0·4)	39·3 (0·5)	44·6 (0·4)	46·3 (0·3)	44·1 (0·4)	41·3 (0·4)	45·2 (0·3)
NCS	45·9 (0·3)	45·6 (0·4)	46·3 (0·3)	45·5 (0·4)	45·8 (0·3)	46·1 (0·3)	42·1 (0·5)	43·9 (0·4)	45·6 (0·3)
stepPlr	43·2 (0·4)	44·6 (0·4)	46·0 (0·3)	41·6 (0·5)	44·5 (0·4)	46·5 (0·3)	34·9 (0·5)	41·7 (0·4)	45·2 (0·4)
rpart	37·9 (0·6)	38·0 (0·6)	45·2 (0·3)	38·2 (0·6)	42·1 (0·4)	44·3 (0·4)	41·5 (0·5)	41·8 (0·5)	43·1 (0·4)
					$n = 500$				
		$p = 50$			$p = 100$			$p = 500$	
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	3·7 (0·1)	9·8 (0·1)	33·9 (0·2)	0·5 (0·0)	3·2 (0·1)	26·7 (0·2)	0·0 (0·0)	0·0 (0·0)	6·7 (0·1)
$\bar{\theta}$ Galton	2·1 (0·0)	2·1 (0·1)	4·6 (0·3)	2·2 (0·1)	2·2 (0·1)	3·6 (0·2)	14·5 (0·4)	2·0 (0·0)	2·6 (0·1)
QCS	3·7 (0·1)	9·8 (0·1)	33·9 (0·2)	0·5 (0·0)	3·2 (0·1)	26·7 (0·2)	0·0 (0·0)	0·0 (0·0)	6·7 (0·1)
$\bar{\theta}$ Skewn.	2·1 (0·0)	2·1 (0·1)	4·6 (0·3)	2·2 (0·1)	2·2 (0·1)	3·6 (0·2)	14·5 (0·4)	2·0 (0·0)	2·6 (0·1)
CC	41·5 (0·2)	44·2 (0·2)	48·0 (0·1)	37·7 (0·2)	42·9 (0·2)	47·8 (0·2)	25·0 (0·2)	34·7 (0·2)	46·6 (0·2)
MC	31·7 (0·2)	38·4 (0·2)	46·7 (0·2)	23·5 (0·2)	33·3 (0·2)	45·4 (0·2)	5·6 (0·1)	16·9 (0·2)	40·5 (0·2)
(Box–Cox) MC	45·9 (0·3)	31·1 (0·2)	38·3 (0·3)	23·4 (0·2)	33·7 (0·2)	45·7 (0·3)	5·8 (0·1)	17·3 (0·2)	40·8 (0·2)
LDA	40·2 (0·2)	42·4 (0·2)	47·4 (0·2)	36·8 (0·3)	40·6 (0·2)	47·4 (0·2)	47·0 (0·2)	47·3 (0·2)	48·0 (0·1)
knn	45·4 (0·2)	47·5 (0·2)	48·3 (0·1)	45·2 (0·3)	47·2 (0·2)	48·2 (0·1)	45·7 (0·2)	47·4 (0·2)	48·3 (0·1)
n-Bayes	33·4 (0·2)	39·1 (0·2)	46·6 (0·2)	33·2 (0·2)	38·4 (0·2)	47·0 (0·2)	28·6 (0·2)	36·9 (0·2)	46·9 (0·2)
SVM	29·7 (0·2)	37·1 (0·2)	46·5 (0·2)	25·2 (0·2)	37·5 (0·2)	47·6 (0·2)	16·4 (0·2)	30·5 (0·2)	47·9 (0·2)
NCS	45·4 (0·2)	46·5 (0·2)	47·7 (0·2)	43·4 (0·3)	45·4 (0·2)	47·7 (0·1)	36·9 (0·2)	40·9 (0·2)	46·9 (0·2)
stepPlr	39·9 (0·2)	42·2 (0·2)	47·4 (0·2)	36·5 (0·2)	40·5 (0·2)	47·3 (0·2)	26·2 (0·2)	35·0 (0·2)	46·6 (0·2)
rpart	22·2 (0·3)	22·4 (0·3)	40·5 (0·3)	20·7 (0·3)	21·0 (0·3)	29·3 (0·4)	22·3 (0·3)	22·4 (0·3)	22·8 (0·4)

Table 5. *Simulation study: independent not identically distributed variables. Misclassification rates multiplied by 100 (with standard errors in brackets; all rounded to one digit after the decimal point) for different methods. Rows 2 and 4 contain the mean of the chosen values of θ in the training sets.*

					$n = 50$				
		$p = 50$			$p = 100$			$p = 500$	
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	24·7 (0·8)	36·1 (0·7)	43·1 (0·5)	19·1 (0·9)	33·0 (0·8)	42·7 (0·5)	5·7 (0·4)	16·8 (0·6)	43·4 (0·5)
$\bar\theta$ Galton	22·1 (2·8)	36·1 (3·1)	56·0 (3·2)	30·1 (3·2)	31·0 (3·0)	59·9 (3·1)	2·2 (0·2)	3·2 (0·6)	21·8 (3·1)
QCS	22·3 (0·7)	32·8 (0·8)	43·4 (0·5)	15·4 (0·8)	27·3 (0·8)	43·6 (0·5)	2·6 (0·3)	11·8 (0·5)	42·9 (0·6)
$\bar\theta$ Skewn.	30·5 (2·9)	37·5 (3·0)	57·2 (2·9)	38·1 (2·7)	50·7 (3·2)	58·2 (3·1)	8·9 (1·8)	9·2 (2·1)	27·0 (3·5)
CC	24·4 (0·7)	35·5 (0·8)	44·0 (0·5)	17·4 (0·5)	29·4 (0·7)	43·2 (0·5)	2·0 (0·2)	12·6 (0·5)	43·7 (0·5)
MC	25·9 (0·6)	36·3 (0·7)	43·6 (0·5)	19·2 (0·7)	31·7 (0·7)	43·5 (0·5)	2·9 (0·2)	14·4 (0·5)	43·7 (0·5)
LDA	40·6 (0·7)	43·4 (0·5)	44·6 (0·5)	31·7 (0·7)	38·0 (0·6)	44·3 (0·4)	25·5 (0·7)	36·0 (0·7)	45·1 (0·4)
knn	39·0 (0·6)	40·3 (0·6)	44·0 (0·4)	29·2 (0·7)	38·2 (0·5)	43·3 (0·5)	12·4 (0·6)	27·3 (0·8)	43·6 (0·5)
n-Bayes	33·8 (0·7)	40·9 (0·6)	44·0 (0·5)	31·6 (0·6)	39·7 (0·6)	43·7 (0·5)	18·8 (0·6)	32·1 (0·7)	44·3 (0·5)
SVM	26·2 (0·6)	35·9 (0·7)	43·1 (0·5)	19·8 (0·6)	33·2 (0·7)	43·9 (0·5)	2·3 (0·2)	13·1 (0·5)	48·5 (0·4)
NCS	32·0 (0·7)	37·5 (0·7)	42·8 (0·6)	24·2 (0·7)	33·1 (0·7)	42·1 (0·6)	7·3 (0·4)	15·8 (0·5)	43·7 (0·4)
stepPlr	28·0 (0·6)	36·3 (0·7)	42·6 (0·5)	20·0 (0·6)	33·6 (0·7)	43·6 (0·5)	2·4 (0·2)	12·9 (0·5)	43·1 (0·5)
rpart	25·5 (1·0)	33·9 (0·9)	39·1 (0·8)	31·3 (0·9)	31·2 (0·9)	39·0 (0·7)	27·7 (0·9)	30·5 (0·9)	42·0 (0·6)
					$n = 100$				
		$p = 50$			$p = 100$			$p = 500$	
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	17·7 (0·4)	30·8 (0·7)	44·2 (0·4)	11·5 (0·4)	24·7 (0·5)	44·7 (0·4)	1·6 (0·2)	10·5 (0·9)	39·5 (0·6)
$\bar\theta$ Galton	13·3 (1·9)	16·9 (2·2)	54·3 (3·2)	14·7 (1·3)	17·5 (2·0)	56·8 (3·2)	18·7 (2·1)	31·7 (2·9)	54·6 (3·2)
QCS	17·2 (0·5)	28·7 (0·6)	44·3 (0·4)	10·7 (0·3)	23·5 (0·6)	43·6 (0·4)	1·2 (0·1)	9·2 (0·9)	38·3 (0·6)
$\bar\theta$ Skewn.	13·0 (1·9)	21·4 (2·7)	53·4 (3·1)	22·9 (1·6)	30·2 (2·6)	53·6 (2·8)	37·2 (1·8)	39·5 (2·0)	61·7 (2·8)
CC	21·1 (0·4)	31·3 (0·6)	44·4 (0·4)	13·2 (0·4)	24·5 (0·5)	42·8 (0·5)	0·7 (0·1)	6·3 (0·3)	35·5 (0·5)
MC	23·9 (0·4)	33·4 (0·5)	44·9 (0·4)	15·8 (0·4)	28·1 (0·4)	43·1 (0·4)	1·4 (0·1)	8·9 (0·3)	37·0 (0·5)
LDA	27·6 (0·5)	36·0 (0·5)	44·6 (0·4)	40·4 (0·6)	43·7 (0·5)	45·6 (0·3)	15·8 (0·4)	27·8 (0·6)	44·3 (0·4)
knn	30·6 (0·5)	41·0 (0·4)	45·9 (0·3)	26·0 (0·5)	35·3 (0·5)	45·2 (0·3)	9·9 (0·4)	27·8 (0·5)	43·0 (0·5)
n-Bayes	28·9 (0·5)	38·2 (0·5)	45·6 (0·3)	25·4 (0·5)	34·4 (0·5)	45·2 (0·4)	13·2 (0·4)	26·8 (0·4)	44·2 (0·4)
SVM	23·0 (0·5)	32·7 (0·5)	44·4 (0·4)	16·0 (0·4)	26·6 (0·4)	42·8 (0·4)	0·9 (0·1)	7·5 (0·3)	34·6 (0·5)
NCS	24·7 (0·5)	32·1 (0·6)	41·6 (0·6)	17·7 (0·4)	25·9 (0·5)	39·5 (0·6)	1·7 (0·1)	6·9 (0·3)	28·7 (0·5)
stepPlr	25·1 (0·5)	36·5 (0·5)	44·4 (0·4)	15·8 (0·4)	28·1 (0·5)	43·7 (0·5)	1·0 (0·1)	7·3 (0·3)	34·2 (0·5)
rpart	18·6 (0·7)	18·5 (0·7)	36·4 (0·6)	17·0 (0·6)	12·7 (0·6)	31·0 (0·7)	13·4 (0·6)	16·4 (0·5)	18·9 (0·6)
					$n = 500$				
		$p = 50$			$p = 100$			$p = 500$	
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	12·1 (0·3)	20·7 (0·4)	40·2 (0·3)	6·6 (0·1)	15·6 (0·1)	37·3 (0·3)	0·1 (0·0)	1·7 (0·1)	49·5 (0·1)
$\bar\theta$ Galton	6·8 (0·7)	7·9 (0·8)	19·9 (2·2)	7·6 (0·5)	8·6 (0·7)	18·2 (1·9)	26·3 (0·8)	16·4 (0·9)	23·5 (1·7)
QCS	9·6 (0·2)	16·8 (0·2)	38·5 (0·3)	5·7 (0·1)	13·4 (0·2)	36·2 (0·3)	0·2 (0·0)	1·3 (0·1)	49·8 (0·0)
$\bar\theta$ Skewn.	2·5 (0·2)	2·2 (0·2)	21·1 (2·5)	9·2 (0·6)	7·4 (0·8)	28·1 (2·4)	30·5 (0·8)	17·4 (0·8)	39·0 (2·1)
CC	17·4 (0·2)	26·2 (0·2)	41·0 (0·2)	9·0 (0·1)	18·3 (0·2)	37·8 (0·2)	0·1 (0·0)	2·1 (0·1)	50·0 (0·0)
MC	20·9 (0·2)	29·3 (0·2)	41·9 (0·2)	12·4 (0·1)	21·8 (0·2)	39·9 (0·2)	0·5 (0·0)	4·2 (0·1)	49·8 (0·1)
LDA	18·8 (0·2)	27·1 (0·2)	41·4 (0·2)	11·3 (0·1)	20·5 (0·2)	38·8 (0·3)	40·8 (0·5)	42·9 (0·4)	48·4 (0·1)
knn	26·6 (0·2)	35·9 (0·3)	47·2 (0·2)	18·8 (0·2)	36·1 (0·3)	47·0 (0·2)	3·4 (0·1)	15·2 (0·3)	48·4 (0·1)
n-Bayes	16·7 (0·2)	26·8 (0·2)	44·0 (0·2)	12·1 (0·1)	22·9 (0·2)	42·7 (0·2)	4·6 (0·1)	14·4 (0·2)	49·0 (0·1)
SVM	17·5 (0·2)	26·5 (0·2)	43·2 (0·2)	9·4 (0·1)	18·8 (0·2)	38·9 (0·3)	0·2 (0·0)	3·1 (0·1)	50·0 (0·0)
NCS	18·6 (0·2)	26·5 (0·2)	39·1 (0·2)	10·2 (0·1)	18·5 (0·2)	35·1 (0·2)	0·2 (0·0)	2·1 (0·1)	50·0 (0·0)
stepPlr	18·5 (0·2)	26·8 (0·2)	41·5 (0·2)	12·1 (0·2)	21·2 (0·2)	39·3 (0·2)	0·2 (0·0)	3·1 (0·1)	50·0 (0·0)
rpart	2·4 (0·1)	6·2 (0·1)	31·9 (0·2)	2·4 (0·1)	3·9 (0·1)	21·6 (0·2)	3·8 (0·1)	3·3 (0·1)	39·6 (0·4)

Table 6. *Simulation study: dependent not identically distributed variables. Misclassification rates multiplied by 100 (with standard errors in brackets; all rounded to one digit after the decimal point) for different methods. Rows 2 and 4 contain the mean of the chosen values of θ in the training sets.*

$n = 50$

	$p = 50$			$p = 100$			$p = 500$		
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	26·2 (0·8)	35·9 (0·8)	43·3 (0·5)	21·1 (1·0)	31·5 (0·7)	43·9 (0·5)	5·8 (0·4)	15·2 (0·6)	40·3 (0·6)
$\bar{\theta}$ Galton	22·1 (2·8)	36·1 (3·1)	56·0 (3·2)	30·1 (3·2)	31·0 (3·0)	59·9 (3·1)	2·2 (0·2)	3·2 (0·6)	21·8 (3·1)
QCS	29·5 (0·9)	36·6 (0·7)	44·3 (0·5)	24·2 (0·8)	35·0 (0·8)	43·9 (0·4)	28·3 (1·2)	28·7 (0·7)	41·2 (0·7)
$\bar{\theta}$ Skewn.	30·5 (2·9)	37·5 (3·0)	57·2 (2·9)	38·1 (2·7)	50·7 (3·2)	58·2 (3·1)	8·9 (1·8)	9·2 (2·1)	27·0 (3·5)
CC	26·9 (0·7)	35·4 (0·7)	42·8 (0·5)	19·6 (0·5)	30·9 (0·7)	44·0 (0·5)	3·3 (0·3)	13·9 (0·5)	39·3 (0·6)
MC	27·5 (0·7)	37·7 (0·5)	42·9 (0·5)	20·3 (0·6)	32·8 (0·7)	42·8 (0·5)	4·0 (0·3)	14·9 (0·5)	39·8 (0·7)
LDA	41·3 (0·6)	42·7 (0·6)	43·8 (0·4)	31·5 (0·7)	39·0 (0·7)	43·5 (0·5)	25·5 (0·8)	34·4 (0·8)	42·6 (0·6)
knn	35·5 (0·7)	39·8 (0·7)	44·1 (0·4)	30·7 (0·6)	40·9 (0·6)	44·8 (0·4)	15·2 (0·7)	28·3 (0·7)	43·9 (0·5)
n-Bayes	34·9 (0·7)	41·4 (0·5)	44·8 (0·5)	31·6 (0·8)	39·4 (0·6)	43·5 (0·5)	20·5 (0·6)	31·3 (0·6)	43·6 (0·5)
SVM	28·3 (0·6)	36·3 (0·7)	43·5 (0·5)	20·9 (0·7)	33·1 (0·7)	43·7 (0·4)	3·7 (0·3)	14·9 (0·5)	39·8 (0·6)
NCS	32·1 (0·7)	36·2 (0·7)	42·7 (0·6)	25·2 (0·6)	31·8 (0·8)	41·7 (0·5)	8·1 (0·4)	16·5 (0·6)	36·0 (0·8)
stepPlr	30·5 (0·7)	36·7 (0·7)	43·5 (0·4)	21·8 (0·7)	32·1 (0·8)	43·7 (0·5)	4·1 (0·3)	14·5 (0·5)	39·4 (0·6)
rpart	31·2 (0·9)	28·3 (1·1)	38·7 (0·7)	30·5 (0·9)	30·3 (0·9)	38·8 (0·8)	29·2 (0·9)	29·9 (0·9)	34·6 (0·9)

$n = 100$

	$p = 50$			$p = 100$			$p = 500$		
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	20·2 (0·6)	29·6 (0·6)	44·6 (0·4)	12·4 (0·4)	23·9 (0·5)	43·9 (0·4)	2·2 (0·4)	11·3 (0·8)	39·2 (0·6)
$\bar{\theta}$ Galton	13·3 (1·9)	16·9 (2·2)	54·3 (3·2)	14·7 (1·3)	17·5 (2·0)	56·8 (3·2)	18·7 (2·1)	31·7 (2·9)	54·6 (3·2)
QCS	23·4 (0·5)	32·3 (0·6)	44·7 (0·3)	15·9 (0·5)	28·2 (0·5)	44·0 (0·4)	4·9 (0·8)	12·2 (0·6)	40·8 (0·6)
$\bar{\theta}$ Skewn.	13·0 (1·9)	21·4 (2·7)	53·4 (3·1)	22·9 (1·6)	30·2 (2·6)	53·6 (2·8)	37·2 (1·8)	39·5 (2·0)	61·7 (2·8)
CC	23·8 (0·4)	32·0 (0·5)	43·9 (0·4)	16·2 (0·4)	25·6 (0·5)	42·9 (0·4)	1·4 (0·1)	8·1 (0·3)	36·2 (0·5)
MC	25·6 (0·4)	33·4 (0·5)	44·7 (0·4)	17·5 (0·4)	28·2 (0·4)	44·3 (0·4)	1·8 (0·1)	10·2 (0·3)	37·9 (0·5)
LDA	31·2 (0·5)	37·0 (0·6)	44·5 (0·4)	41·5 (0·5)	44·0 (0·4)	45·6 (0·3)	14·8 (0·4)	27·0 (0·5)	43·3 (0·5)
knn	32·8 (0·6)	39·1 (0·5)	45·8 (0·3)	25·1 (0·5)	37·9 (0·5)	44·9 (0·4)	11·2 (0·5)	26·8 (0·5)	44·4 (0·4)
n-Bayes	30·0 (0·5)	38·7 (0·6)	46·1 (0·3)	25·3 (0·5)	35·5 (0·5)	44·8 (0·4)	14·2 (0·4)	26·6 (0·4)	44·5 (0·4)
SVM	24·7 (0·5)	32·8 (0·5)	44·0 (0·4)	16·6 (0·4)	27·1 (0·5)	43·2 (0·4)	1·6 (0·1)	9·0 (0·3)	37·0 (0·5)
NCS	26·4 (0·5)	32·1 (0·5)	42·4 (0·5)	18·6 (0·4)	25·6 (0·5)	38·9 (0·5)	2·2 (0·2)	8·2 (0·3)	28·4 (0·6)
stepPlr	27·6 (0·5)	34·1 (0·5)	44·5 (0·5)	19·9 (0·4)	28·3 (0·5)	43·4 (0·4)	1·9 (0·1)	9·2 (0·3)	36·6 (0·5)
rpart	12·8 (0·6)	18·5 (0·7)	35·6 (0·6)	16·4 (0·6)	13·5 (0·6)	30·1 (0·7)	22·8 (0·6)	16·1 (0·5)	16·9 (0·6)

$n = 500$

	$p = 50$			$p = 100$			$p = 500$		
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	13·0 (0·4)	20·4 (0·4)	40·1 (0·3)	7·6 (0·1)	15·7 (0·2)	36·9 (0·3)	0·2 (0·0)	1·9 (0·1)	24·9 (0·3)
$\bar{\theta}$ Galton	6·8 (0·7)	7·9 (0·8)	19·9 (2·2)	7·6 (0·5)	8·6 (0·7)	18·2 (1·9)	26·3 (0·8)	16·4 (0·9)	23·5 (1·7)
QCS	14·8 (0·2)	21·7 (0·2)	40·6 (0·3)	9·6 (0·2)	19·6 (0·2)	38·9 (0·3)	0·4 (0·0)	2·9 (0·1)	28·6 (0·3)
$\bar{\theta}$ Skewn.	2·5 (0·2)	2·2 (0·2)	21·1 (2·5)	9·2 (0·6)	7·4 (0·8)	28·1 (2·4)	30·5 (0·8)	17·4 (0·8)	39·0 (2·1)
CC	20·6 (0·2)	27·5 (0·2)	41·2 (0·2)	12·0 (0·1)	19·9 (0·2)	38·0 (0·2)	0·4 (0·0)	3·0 (0·1)	24·9 (0·2)
MC	22·8 (0·2)	29·6 (0·2)	42·3 (0·3)	14·3 (0·2)	22·7 (0·2)	39·7 (0·2)	0·8 (0·0)	4·8 (0·1)	28·5 (0·2)
LDA	21·5 (0·2)	27·9 (0·2)	41·2 (0·2)	14·1 (0·1)	22·4 (0·2)	38·6 (0·2)	41·7 (0·4)	42·8 (0·4)	47·4 (0·2)
knn	27·9 (0·3)	36·4 (0·2)	47·3 (0·2)	22·5 (0·3)	33·0 (0·3)	46·6 (0·2)	5·4 (0·2)	17·2 (0·3)	43·1 (0·2)
n-Bayes	17·6 (0·2)	27·6 (0·2)	44·1 (0·2)	13·0 (0·2)	23·1 (0·2)	42·3 (0·2)	4·8 (0·1)	14·6 (0·1)	38·8 (0·2)
SVM	20·5 (0·2)	27·2 (0·2)	43·8 (0·2)	11·9 (0·1)	21·0 (0·2)	38·3 (0·2)	0·5 (0·0)	3·7 (0·1)	27·1 (0·3)
NCS	20·7 (0·2)	27·2 (0·2)	39·1 (0·2)	11·8 (0·1)	19·7 (0·2)	34·8 (0·2)	0·4 (0·0)	2·7 (0·1)	18·9 (0·2)
stepPlr	21·9 (0·2)	28·0 (0·2)	41·3 (0·2)	14·8 (0·2)	22·8 (0·2)	38·8 (0·2)	0·6 (0·0)	4·2 (0·1)	28·2 (0·3)
rpart	2·8 (0·1)	6·8 (0·1)	31·8 (0·2)	2·4 (0·1)	2·9 (0·1)	21·6 (0·2)	3·9 (0·1)	3·1 (0·1)	2·8 (0·1)

Table 7. *Simulation study: Beta distributed variables differing between balanced classes. Misclassification rates multiplied by 100 (with standard errors in brackets; all rounded to one digit after the decimal point) for different methods. Rows 2 and 4 contain the mean of the chosen values of θ in the training sets.*

	$p = 50$			$p = 100$			$p = 500$		
$n = 50$	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	10·1 (0·7)	23·8 (0·9)	41·9 (0·7)	3·6 (0·6)	14·3 (0·9)	40·4 (0·7)	0·0 (0·0)	0·2 (0·1)	22·6 (0·8)
$\bar{\theta}$ Galton	22·7 (3·7)	26·0 (3·7)	49·7 (3·8)	14·9 (3·1)	18·9 (3·3)	46·1 (3·9)	2·0 (0·0)	2·3 (0·3)	8·7 (2·3)
QCS	7·2 (0·6)	19·3 (1·0)	40·8 (0·7)	3·0 (0·5)	10·5 (0·8)	38·8 (0·8)	0·0 (0·0)	0·2 (0·1)	20·8 (0·9)
$\bar{\theta}$ Skewn.	8·7 (2·1)	12·8 (2·5)	31·1 (3·5)	13·9 (3·0)	10·7 (2·5)	35·0 (3·8)	2·0 (0·0)	3·2 (1·0)	13·3 (3·0)
CC	45·1 (0·4)	44·3 (0·5)	44·2 (0·4)	44·6 (0·4)	44·7 (0·4)	43·8 (0·4)	44·3 (0·4)	44·0 (0·4)	44·2 (0·5)
MC	31·8 (0·9)	37·9 (0·7)	44·2 (0·5)	24·3 (0·6)	35·0 (0·7)	43·6 (0·5)	6·5 (0·3)	20·0 (0·6)	41·6 (0·5)
LDA	44·3 (0·4)	45·0 (0·4)	44·0 (0·4)	44·6 (0·4)	44·8 (0·4)	44·0 (0·4)	45·2 (0·4)	44·9 (0·4)	44·9 (0·4)
knn	35·5 (0·9)	43·2 (0·5)	44·8 (0·4)	35·6 (0·7)	41·7 (0·6)	44·6 (0·4)	36·9 (0·7)	42·7 (0·5)	44·3 (0·4)
n-Bayes	6·3 (0·5)	14·4 (0·7)	38·2 (0·8)	3·0 (0·3)	10·1 (0·6)	35·8 (0·8)	0·0 (0·0)	1·1 (0·2)	25·6 (0·7)
SVM	43·5 (0·5)	44·0 (0·4)	44·0 (0·5)	44·8 (0·5)	44·1 (0·4)	44·1 (0·4)	44·3 (0·4)	43·6 (0·4)	44·6 (0·4)
NCS	43·7 (0·5)	43·8 (0·5)	44·7 (0·4)	44·2 (0·5)	43·5 (0·5)	44·7 (0·4)	44·7 (0·4)	44·1 (0·4)	44·6 (0·4)
stepPlr	45·0 (0·4)	44·1 (0·4)	44·0 (0·4)	44·4 (0·4)	44·0 (0·5)	44·0 (0·4)	44·0 (0·5)	43·6 (0·5)	44·7 (0·4)
rpart	23·8 (1·0)	21·7 (1·1)	36·0 (1·2)	18·3 (0·9)	23·4 (1·1)	33·1 (1·3)	16·5 (0·8)	18·6 (0·9)	24·9 (1·2)
$n = 100$	$p = 50$			$p = 100$			$p = 500$		
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	5·1 (0·4)	16·9 (0·8)	39·3 (0·8)	1·2 (0·2)	8·0 (0·6)	36·9 (0·8)	0·0 (0·0)	0·7 (0·2)	19·6 (1·0)
$\bar{\theta}$ Galton	5·4 (1·3)	14·3 (2·9)	37·7 (4·0)	5·6 (1·6)	8·7 (2·0)	36·5 (4·0)	4·9 (1·6)	11·9 (2·8)	27·9 (4·2)
QCS	3·7 (0·3)	15·2 (0·7)	38·4 (0·7)	0·6 (0·2)	4·9 (0·3)	34·1 (0·8)	0·0 (0·0)	1·3 (0·3)	18·7 (1·3)
$\bar{\theta}$ Skewn.	3·2 (0·3)	5·4 (1·0)	25·5 (3·3)	3·4 (1·0)	4·1 (0·9)	21·0 (3·3)	9·7 (2·6)	17·2 (3·3)	26·8 (3·9)
CC	46·3 (0·3)	45·9 (0·3)	46·4 (0·3)	45·7 (0·3)	46·1 (0·3)	46·3 (0·3)	45·3 (0·3)	46·2 (0·3)	45·9 (0·3)
MC	28·6 (0·6)	38·0 (0·6)	44·8 (0·4)	20·7 (0·5)	31·0 (0·6)	44·3 (0·4)	3·7 (0·2)	15·3 (0·4)	41·3 (0·5)
LDA	45·9 (0·3)	46·5 (0·2)	46·5 (0·3)	45·6 (0·4)	46·1 (0·3)	45·9 (0·3)	46·4 (0·3)	46·0 (0·3)	46·5 (0·3)
knn	33·8 (0·7)	39·7 (0·7)	45·5 (0·3)	32·0 (0·6)	42·7 (0·5)	45·7 (0·4)	34·5 (0·6)	42·6 (0·5)	45·5 (0·3)
n-Bayes	2·7 (0·2)	9·0 (0·5)	34·5 (0·8)	1·6 (0·2)	5·3 (0·4)	30·1 (0·7)	0·0 (0·0)	0·3 (0·1)	18·8 (0·6)
SVM	42·4 (0·5)	45·7 (0·4)	46·1 (0·3)	39·4 (0·6)	46·1 (0·3)	45·8 (0·3)	45·4 (0·3)	45·8 (0·3)	46·4 (0·3)
NCS	45·0 (0·4)	45·8 (0·4)	46·3 (0·3)	44·7 (0·4)	46·0 (0·3)	45·8 (0·3)	45·8 (0·3)	45·6 (0·3)	46·5 (0·3)
stepPlr	46·0 (0·3)	46·5 (0·3)	45·6 (0·3)	45·6 (0·3)	46·0 (0·3)	45·9 (0·3)	46·0 (0·3)	45·7 (0·3)	45·7 (0·3)
rpart	18·3 (0·7)	17·7 (0·6)	32·6 (1·1)	10·9 (0·5)	14·7 (0·5)	24·7 (1·0)	12·0 (0·5)	9·9 (0·5)	18·4 (0·8)
$n = 500$	$p = 50$			$p = 100$			$p = 500$		
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	2·2 (0·1)	10·4 (0·4)	35·1 (0·7)	0·2 (0·0)	2·9 (0·2)	27·1 (0·7)	0·0 (0·0)	0·0 (0·0)	5·3 (0·2)
$\bar{\theta}$ Galton	2·6 (0·1)	3·2 (0·2)	20·0 (3·1)	2·5 (0·2)	2·8 (0·2)	7·2 (1·6)	27·0 (4·2)	2·9 (0·9)	3·2 (0·2)
QCS	1·9 (0·1)	9·6 (0·4)	34·8 (0·7)	0·2 (0·0)	2·6 (0·2)	26·0 (0·7)	0·0 (0·0)	0·0 (0·0)	4·4 (0·2)
$\bar{\theta}$ Skewn.	2·6 (0·1)	2·9 (0·1)	17·0 (2·9)	2·3 (0·1)	2·4 (0·1)	4·7 (0·5)	31·8 (4·5)	2·0 (0·0)	2·9 (0·2)
CC	48·2 (0·1)	48·0 (0·1)	48·2 (0·1)	48·2 (0·1)	48·1 (0·1)	48·3 (0·1)	47·9 (0·1)	48·2 (0·1)	48·2 (0·1)
MC	24·6 (0·5)	33·9 (0·5)	44·1 (0·4)	16·8 (0·4)	26·7 (0·5)	42·3 (0·5)	1·3 (0·1)	7·5 (0·2)	33·2 (0·4)
LDA	48·5 (0·1)	48·1 (0·1)	48·1 (0·1)	48·2 (0·1)	48·1 (0·1)	48·1 (0·1)	48·1 (0·1)	48·2 (0·1)	48·1 (0·1)
knn	23·6 (0·7)	34·9 (0·7)	45·4 (0·4)	23·7 (0·6)	34·6 (0·6)	46·8 (0·3)	24·3 (0·9)	35·8 (0·4)	46·9 (0·2)
n-Bayes	0·7 (0·1)	3·6 (0·2)	27·6 (0·9)	0·2 (0·0)	1·3 (0·1)	19·5 (0·7)	0·0 (0·0)	0·0 (0·0)	5·1 (0·2)
SVM	18·1 (0·4)	30·6 (0·5)	48·3 (0·1)	26·9 (0·4)	37·2 (0·4)	48·0 (0·1)	45·5 (0·3)	47·8 (0·2)	48·1 (0·1)
NCS	46·8 (0·4)	47·3 (0·3)	48·6 (0·1)	47·1 (0·3)	47·6 (0·2)	48·2 (0·1)	48·2 (0·1)	48·0 (0·1)	48·1 (0·1)
stepPlr	48·3 (0·1)	48·0 (0·1)	48·4 (0·1)	48·0 (0·1)	48·2 (0·1)	48·5 (0·1)	48·2 (0·1)	48·2 (0·1)	47·9 (0·1)
rpart	5·1 (0·2)	7·7 (0·4)	27·2 (1·0)	3·4 (0·1)	5·2 (0·2)	16·3 (0·8)	2·7 (0·1)	2·7 (0·1)	5·6 (0·2)

Table 8. *Simulation study: Beta distributed variables differing between unbalanced classes. Misclassification rates multiplied by 100 (with standard errors in brackets; all rounded to one digit after the decimal point) for different methods. Rows 2 and 4 contain the mean of the chosen values of θ in the training sets.*

					$n = 50$				
		$p = 50$			$p = 100$			$p = 500$	
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	10·9 (0·6)	19·6 (0·6)	25·0 (0·4)	6·5 (0·5)	15·8 (0·5)	23·4 (0·2)	0·9 (0·2)	13·8 (0·5)	23·8 (0·1)
$\bar{\theta}$ Galton	22·7 (3·7)	34·7 (4·1)	40·9 (4·1)	18·8 (3·5)	29·1 (4·2)	36·0 (3·9)	2·0 (0·0)	2·0 (0·0)	2·2 (0·2)
QCS	8·9 (0·6)	17·6 (0·7)	24·3 (0·4)	5·6 (0·5)	14·4 (0·6)	22·9 (0·3)	0·4 (0·1)	11·7 (0·5)	23·7 (0·1)
$\bar{\theta}$ Skewn.	9·0 (2·0)	24·4 (3·6)	35·1 (4·0)	17·3 (3·3)	18·7 (3·4)	31·4 (4·1)	2·0 (0·0)	2·2 (0·2)	3·1 (1·0)
CC	30·8 (0·6)	31·6 (0·7)	30·4 (0·6)	25·6 (0·4)	26·8 (0·5)	26·3 (0·4)	23·6 (0·1)	23·6 (0·1)	23·6 (0·1)
MC	24·4 (0·6)	28·3 (0·6)	29·4 (0·6)	19·3 (0·5)	23·5 (0·4)	25·5 (0·3)	13·0 (0·5)	20·1 (0·3)	23·3 (0·1)
LDA	39·7 (0·6)	41·0 (0·5)	39·2 (0·6)	25·1 (0·3)	25·4 (0·4)	25·6 (0·3)	24·0 (0·0)	24·0 (0·0)	24·0 (0·0)
knn	24·4 (0·5)	26·6 (0·5)	27·2 (0·5)	26·1 (0·5)	25·4 (0·5)	26·7 (0·5)	23·6 (0·4)	24·1 (0·3)	24·6 (0·3)
n-Bayes	6·6 (0·4)	14·6 (0·5)	23·1 (0·3)	4·1 (0·3)	13·2 (0·5)	23·1 (0·2)	0·5 (0·1)	9·1 (0·5)	23·0 (0·2)
SVM	23·1 (0·2)	24·0 (0·0)	24·0 (0·0)	24·0 (0·0)	24·0 (0·0)	24·0 (0·0)	24·0 (0·0)	24·0 (0·0)	24·0 (0·0)
NCS	24·0 (0·0)	24·0 (0·0)	24·0 (0·0)	24·0 (0·0)	24·0 (0·0)	24·0 (0·0)	23·9 (0·1)	23·9 (0·1)	24·0 (0·0)
stepPlr	23·5 (0·1)	24·0 (0·0)	24·0 (0·0)	24·0 (0·0)	24·0 (0·0)	24·0 (0·0)	23·7 (0·1)	24·0 (0·0)	24·0 (0·0)
rpart	19·1 (0·8)	19·1 (0·8)	24·8 (0·7)	16·5 (0·6)	15·7 (0·6)	21·4 (0·7)	11·2 (0·5)	11·9 (0·5)	19·0 (0·6)
					$n = 100$				
		$p = 50$			$p = 100$			$p = 500$	
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	6·0 (0·4)	17·6 (0·6)	30·1 (0·5)	1·7 (0·3)	9·2 (0·6)	25·8 (0·4)	0·0 (0·0)	1·0 (0·2)	19·6 (0·4)
$\bar{\theta}$ Galton	12·9 (2·8)	21·4 (3·3)	42·6 (3·7)	14·3 (3·1)	18·3 (3·3)	42·3 (3·8)	2·0 (0·0)	6·0 (1·7)	30·1 (4·1)
QCS	4·4 (0·4)	14·0 (0·6)	28·5 (0·5)	0·8 (0·1)	5·9 (0·4)	24·8 (0·4)	0·0 (0·0)	1·4 (0·3)	19·9 (0·4)
$\bar{\theta}$ Skewn.	5·3 (1·1)	9·0 (1·9)	37·7 (3·7)	6·0 (1·9)	4·9 (1·2)	30·0 (3·5)	3·9 (1·4)	12·2 (2·9)	33·5 (4·1)
CC	34·6 (0·5)	35·0 (0·6)	36·0 (0·6)	30·1 (0·5)	31·2 (0·5)	31·8 (0·5)	25·0 (0·0)	24·9 (0·0)	25·1 (0·1)
MC	24·5 (0·5)	29·1 (0·5)	33·9 (0·6)	17·3 (0·4)	25·2 (0·4)	28·3 (0·4)	6·0 (0·3)	15·6 (0·4)	24·7 (0·1)
LDA	29·2 (0·4)	29·9 (0·5)	29·8 (0·5)	42·5 (0·4)	43·1 (0·4)	42·8 (0·4)	24·8 (0·0)	24·8 (0·0)	24·9 (0·0)
knn	24·0 (0·4)	24·7 (0·2)	26·9 (0·4)	23·6 (0·4)	28·5 (0·6)	28·1 (0·5)	23·9 (0·5)	24·8 (0·1)	28·1 (0·5)
n-Bayes	3·0 (0·2)	8·8 (0·4)	22·8 (0·3)	1·6 (0·2)	7·1 (0·3)	22·5 (0·3)	0·0 (0·0)	1·4 (0·1)	21·0 (0·3)
SVM	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)
NCS	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)
stepPlr	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	24·9 (0·0)	25·0 (0·0)	25·0 (0·0)	24·8 (0·0)	24·9 (0·0)	24·9 (0·0)
rpart	11·2 (0·5)	12·8 (0·5)	19·1 (0·6)	9·1 (0·4)	13·4 (0·5)	17·9 (0·5)	7·9 (0·3)	10·1 (0·4)	13·9 (0·5)
					$n = 500$				
		$p = 50$			$p = 100$			$p = 500$	
	100%	50%	10%	100%	50%	10%	100%	50%	10%
QCG	2·3 (0·1)	10·2 (0·5)	31·5 (0·7)	0·3 (0·0)	2·9 (0·2)	26·6 (0·6)	0·0 (0·0)	0·0 (0·0)	5·4 (0·2)
$\bar{\theta}$ Galton	2·7 (0·2)	4·7 (1·0)	24·1 (3·2)	2·5 (0·1)	2·9 (0·2)	18·2 (2·9)	36·6 (4·6)	3·9 (1·4)	3·7 (0·2)
QCS	1·9 (0·1)	9·5 (0·4)	31·6 (0·8)	0·2 (0·0)	2·5 (0·2)	25·0 (0·6)	0·0 (0·0)	0·0 (0·0)	4·6 (0·2)
$\bar{\theta}$ Skewn.	2·7 (0·2)	3·3 (0·2)	20·3 (2·8)	2·3 (0·1)	3·1 (0·2)	12·6 (2·1)	29·8 (4·4)	2·0 (0·0)	3·5 (0·2)
CC	42·6 (0·3)	43·3 (0·3)	43·3 (0·3)	40·9 (0·3)	41·0 (0·3)	41·2 (0·3)	29·6 (0·3)	30·5 (0·3)	31·3 (0·3)
MC	22·7 (0·4)	31·4 (0·4)	40·5 (0·4)	15·3 (0·4)	24·2 (0·4)	37·4 (0·3)	1·3 (0·1)	7·3 (0·2)	26·2 (0·2)
LDA	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	46·6 (0·2)	46·8 (0·2)	46·9 (0·2)
knn	19·3 (0·4)	24·5 (0·2)	25·0 (0·0)	19·5 (0·4)	24·3 (0·3)	26·3 (0·2)	18·8 (0·3)	24·6 (0·2)	25·2 (0·1)
n-Bayes	0·7 (0·1)	3·6 (0·2)	19·8 (0·5)	0·3 (0·0)	1·3 (0·1)	15·9 (0·5)	0·0 (0·0)	0·0 (0·0)	8·9 (0·3)
SVM	14·9 (0·3)	25·0 (0·0)	25·0 (0·0)	19·4 (0·3)	25·0 (0·0)	25·0 (0·0)	22·3 (0·2)	24·9 (0·0)	25·0 (0·0)
NCS	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)	25·0 (0·0)
stepPlr	24·9 (0·0)	25·0 (0·0)	25·0 (0·0)	24·9 (0·0)	24·9 (0·0)	25·0 (0·0)	25·7 (0·1)	25·0 (0·0)	25·0 (0·0)
rpart	3·6 (0·2)	9·8 (0·4)	17·6 (0·7)	2·3 (0·1)	3·7 (0·2)	13·2 (0·5)	2·0 (0·1)	1·9 (0·1)	4·4 (0·2)