



Y_p : RNA plasma, Y_d : RNA DBS, X : covs
 1 = detected, 0 = not

sens: $P(Y_d=1 | Y_p=1, X) = \gamma_1$ non-differential indep of X, not edge X and Y_d

spec: $P(Y_d=0 | Y_p=0, X) = 1 - \gamma_0$ $\gamma_0 = P(Y_d=1 | Y_p=0, X)$ non-differential

Our model: logistic or RR model for X and Y_p

Y_p not observed but misclassification probs "known" for generalized linear model theory and Newhaus (1999)

$\mu_p(X) = P(Y_p=1 | X)$ follows a GLM

$g_p \{ \mu_p(X) \} = \eta_p = X\beta_p$ $Y_p \sim \text{binomial}$
 $g_p = \text{logit, logistics reg}$; $g_p = \text{log, RR reg}$

Newhaus '99

Note: $\mu(X) = P(Y_d=1 | X) = \gamma_0 + (\gamma_1 - \gamma_0)\mu_p(X)$
 which is a GLM where link g is

$$g \{ \mu(X) \} = g_p \left\{ \frac{\mu_p(X) - \gamma_0}{\gamma_1 - \gamma_0} \right\} = \eta_p$$

$$\mu(X) = g_p^{-1} \{ \eta_p \} (\gamma_1 - \gamma_0) + \gamma_0$$

g_p : logit for logistic g_p : log for RR

This is a GLM with a modified link function

Stata allows you to write these. I wrote one for ORs and one for RRs.

Need to calculate

$$\frac{\partial \mu}{\partial \eta} \quad \begin{matrix} \text{logit} \\ (\delta_1 - \delta_0) \exp(\eta) (1 + \exp(\eta)) \end{matrix} \quad \begin{matrix} \text{log} \\ (\delta_1 - \delta_0) e^\eta \end{matrix}$$

$$\frac{\partial^2 \mu}{\partial \eta^2} \quad (\delta_1 - \delta_0) \frac{e^\eta (1 - e^\eta)}{(1 + e^\eta)^3} \quad (\delta_1 - \delta_0) e^\eta$$

These are programmed in two ado file

"mislogit" and "mislog"

From Ingrid: $\delta_1 = 0.808$, $1 - \delta_0 = 0.873$

$$\Rightarrow \delta_0 = 0.127$$

$$\delta_1 - \delta_0 = 0.681$$

type "sysdir" in Stata. It tells you where "ado" files go. For me its

~/Library/Application Support/Stata/ado/personal

If you put the ado files there, the code works

```

program mislog
version 14.2
args todo eta mu return
if `todo' == -1 {
    /* Set global macros for output */
    global SGLM_It "misclassified log"
    global SGLM_If "log( (p-0.127)/0.681)"
    exit
}
if `todo' == 0 {
    /* set eta =g(mu) */
    generate double `eta' = log( (`mu' - 0.127)/0.681 )
    exit
}
if `todo' == 1 {
    /* set g^{-1}(eta) */
    generate double `mu' = exp(`eta')*0.681 + 0.127
    exit
}
if `todo' == 2 {
    /* set return = dmu/deta */
    local d2 = 0.681*exp(`eta')
    generate double `return' = 0.681*exp(`eta')
    exit
}
if `todo' == 3 {
    /* set return =d2mu/dmu2 */
    generate double `return' = 0.681*exp(`eta')
    exit
}
display as error "Unknown call to glm link function"
exit 198
end

```

```

program mislogit
version 14.2
args todo eta mu return
if `todo' == -1 {
    /* Set global macros for output */
    global SGLM_It "logit with missclassification"
    global SGLM_If "logit(((p - 0.127)/0.681))"
    exit
}

```

```

}
if `todo' == 0 {
    /* set eta =g(mu) */
    generate double `eta' = logit( ((`mu' - 0.127)/0.681) )
    exit
}
if `todo' == 1 {
    /* set g^{-1}(eta) */
    generate double `mu' = (invlogit(`eta')*0.681) + 0.127
    exit
}
if `todo' == 2 {
    /* set return = dmu/deta */
    generate double `return' = 0.681*invlogit(`eta')*(1-invlogit(`eta'))
    exit
}
if `todo' == 3 {
    /* set return =d2mu/dmu2 */
    generate double `return' = 0.681*( exp(`eta')*(1-exp(`eta')) )/( (1+exp(`eta'))^3 )
    exit
}
display as error "Unknown call to glm link function"
exit 198
end

```