

Here the predictions from the models coincide nearly perfectly in the region where most of the data are concentrated and are very similar elsewhere. It is thus futile to attempt to empirically distinguish between the logit and probit links unless one has a large sample.

6.3 Which treatment is best for toenail infection?

In the previous section, we described conventional modeling of dichotomous responses where it is assumed that the responses are conditionally independent given the covariates x_i . We are now ready to consider multilevel models for *clustered* dichotomous responses, which are dependent even after conditioning on the covariates.

Lesaffre and Spiessens (2001) analyzed data provided by De Backer et al. (1998) from a randomized, double-blind trial of treatments for toenail infection (dermatophyte onychomycosis). Toenail infection is common with a prevalence of about 2% to 3% in the United States, and a much higher prevalence among diabetics and the elderly. The infection is caused by a fungus and does not only disfigure the nails but can also cause physical pain and impair the ability to work.

In this clinical trial, 378 patients were randomly allocated into two oral antifungal treatments (250 mg/day terbinafine and 200 mg/day itraconazole) and evaluated at seven visits, at weeks 0, 4, 8, 12, 24, 36, and 48. One outcome is onycholysis, the degree of separation of the nail plate from the nail bed, which has been dichotomized ("moderate or severe" versus "none or mild") and is available for 294 patients.

The dataset `toenail.dta` contains the following variables:

- `patient`: patient identifier
- `outcome`: onycholysis (separation of nail plate from nail bed) (0: none or mild; 1: moderate or severe)
- `treatment`: treatment group (0: itraconazole; 1: terbinafine)
- `visit`: visit number (1, 2, ..., 7)
- `month`: exact timing of visit in months

The main research question is whether the treatments differ in their efficacy. In other words, do patients receiving one treatment experience a greater decrease in their probability of having onycholysis than those receiving the other treatment?

6.4 Longitudinal data structure

Before investigating the research question, we look at the longitudinal structure of the toenail data. We can use the `xtdescribe` command introduced in chapter 5 because the data were intended to be balanced with seven visits planned for the same set of weeks for each patient (although the exact timing of the visits varied between patients).

use `http://www`
`xtdescribe if`
`patient: 1, 2`
`visit: 1, 2`
`Delt:`
`(pat`
 Distribution of

Freq.	Per
224	7
21	
10	
6	
5	
5	
4	
3	
3	
13	
294	1

The dataset is not cally, 224 patients 6th visit ("11111. patients dropped c pattern is sometin missingness, which

As discussed i incomplete data s who attended all the model is corre responses are mis

6.5 Populatio

A useful graphica with onycholysis be used to produ

```
. label def
. label val
. graph bar
> ytitle(Pr
```

Here we defined

```

. use http://www.stata-press.com/data/mlmus2/toenail, clear
. xtdescribe if outcome < ., i(patient) t(visit)
patient: 1, 2, ..., 383
visit: 1, 2, ..., 7
Delta(visit) = 1; (7-1)+1 = 7
(patient*visit uniquely identifies each observation)

```

Distribution of T_i:			min	5%	25%	50%	75%	95%	max
			1	3	7	7	7	7	7
Freq.	Percent	Cum.	Pattern						
224	76.19	76.19	1111111						
21	7.14	83.33	11111.1						
10	3.40	86.73	1111.11						
6	2.04	88.78	111....						
5	1.70	90.48	1.....						
5	1.70	92.18	11111..						
4	1.36	93.54	11111..						
3	1.02	94.56	11.....						
3	1.02	95.58	111.111						
13	4.42	100.00	(other patterns)						
294	100.00		XXXXXXX						

The dataset is not balanced since all patients did not attend all planned visits. Specifically, 224 patients have complete data (the pattern "1111111"), 21 patients missed the 6th visit ("11111.1"), 10 patients missed the 5th visit ("1111.11"), and most other patients dropped out at some point, never returning after missing a visit. The latter pattern is sometimes referred to as *monotone missingness* in contrast to *intermittent missingness*, which follows no particular pattern.

As discussed in section 5.9, a nice feature of maximum likelihood estimation for incomplete data such as these is that all information is used. Thus not only patients who attended all visits, but also patients with missing visits contribute information. If the model is correctly specified, maximum likelihood estimates are consistent when the responses are missing at random (MAR).

6.5 Population-averaged or marginal probabilities

A useful graphical display of the data is a bar plot showing the proportion of patients with onycholysis at each visit by treatment group. The following Stata commands can be used to produce the graph shown in figure 6.6:

```

label define tr 0 "Itraconazole" 1 "Terbinafine"
label values treatment tr
graph bar (mean) proportion= outcome, over(visit) by(treatment)
title(Proportion with onycholysis)

```

Use the defined value labels for treatment to make them appear on the graph.

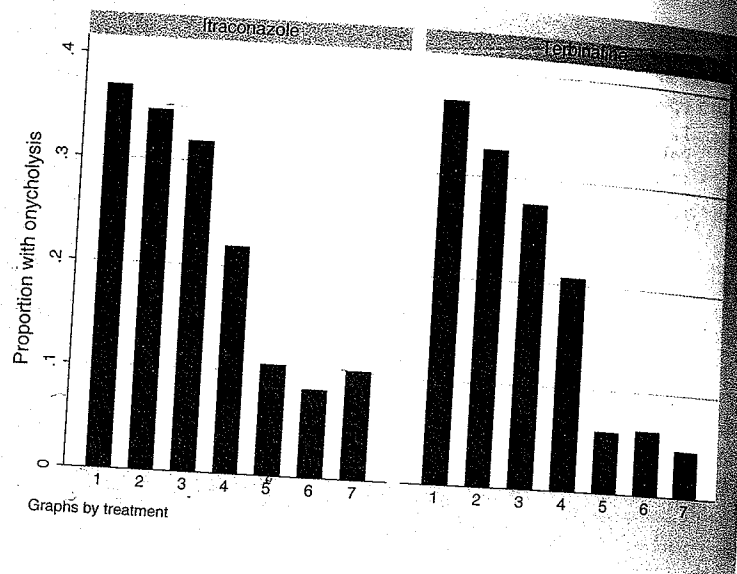


Figure 6.6: Bar plot of proportion of patients with toenail infection by visit and treatment group

We used the visit number `visit` to define the bars instead of the exact timing of the visit month because there would generally not be enough patients with the same timing to reliably estimate the proportions. An alternative display is a line graph, plotting the observed proportions at each visit against time. For this graph, it is better to use the average time associated with each visit for the x axis than using visit number since the visits were not equally spaced. Both the proportions and average times for each visit in each treatment group can be obtained by using the `egen` command with the `mean()` function:

```
. egen prop = mean(outcome), by(treatment visit)
. egen mn_month = mean(month), by(treatment visit)
. twoway line prop mn_month, by(treatment) sort
> xtitle(Time in months) ytitle(Proportion of onycholysis)
```

The resulting graph is shown in figure 6.7.

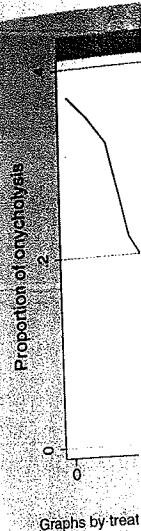


Figure 6.7: Line plot of proportion of patients with toenail infection by visit and treatment group

The proportions of onycholysis are the probabilities of onycholysis for each treatment group. We are interested in the differences in these probabilities, which may vary with the covariates.

Instead of estimating the probabilities, we can attempt to estimate the log-odds of onycholysis. We then do not need to know the baseline probabilities but can directly use the results from a logistic regression model. This model for the data is

$$\text{logit}\{$$

where x_{2j} represents the vector of covariates for containing both the baseline β_2 , and the difference in the log-odds between the itraconazole and terbinafine treatment groups. can be viewed as the difference in the log-odds between the two treatment groups. The model makes the usual assumption of conditional independence of the outcomes given the covariates.

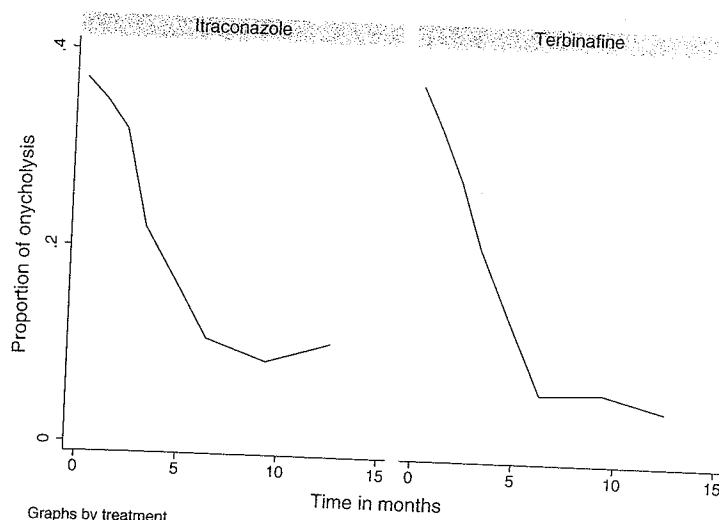


Figure 6.7: Line plot of proportion of patients with toenail infection by visit and treatment group

The proportions shown in figure 6.7 represent the estimated average (or marginal) probabilities of onycholysis given the two covariates: time since randomization and treatment group. We are not attempting to estimate individual patients' personal probabilities, which may vary substantially, but are considering the population averages, given the covariates.

Instead of estimating the probabilities for each combination of visit and treatment, we can attempt to obtain smooth curves of the estimated probability as a function of time. We then no longer have to group observations for the same visit number together but can directly use the exact timing of the visits. One way to accomplish this is by using a logistic regression model with month, treatment, and their interaction as covariates. This model for the dichotomous outcome y_{ij} at visit i for patient j can be written as

$$\text{logit}\{\Pr(y_{ij}=1|\mathbf{x}_{ij})\} = \beta_1 + \beta_2 x_{2j} + \beta_3 x_{3ij} + \beta_4 x_{2j} x_{3ij} \quad (6.5)$$

where x_{2j} represents treatment, x_{3ij} represents month, and $\mathbf{x}_{ij} = (x_{2j}, x_{3ij})'$ is a vector containing both covariates. This model allows for a difference between groups at baseline β_2 , and linear changes in the log odds of onycholysis over time with slope β_3 in the itraconazole group and slope $\beta_3 + \beta_4$ in the terbinafine group. Therefore, β_4 , the difference in the rate of improvement (on the log odds scale) between treatment groups, can be viewed as the treatment effect (terbinafine versus itraconazole). This model makes the unrealistic assumption that the responses for a given patient are conditionally independent after controlling for the included covariates. We will relax this assumption in the next section.

We now check how well predicted probabilities from the logistic regression model correspond to the observed proportions in figure 6.7. The predicted probabilities are obtained as follows:

```
. generate trt_month = treatment*month
. logit outcome treatment month trt_month, or
Logistic regression
```

Log likelihood = -908.00747

```
Number of obs   =    1908
LR chi2(3)      =    164.47
Prob > chi2     =    0.0000
Pseudo R2      =    0.0830
```

outcome	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
treatment	.9994185	.1560558	-0.00	0.997	.7359277	1.357249
month	.8434052	.0199212	-7.21	0.000	.8052504	.8833679
trt_month	.934988	.0350845	-1.79	0.073	.8686913	1.006344

```
. predict prob, p
```

Plotting these together with the observed proportions using the command

```
. twoway (line prop mn_month, sort) (line prob month, sort lpatt(dash)),
> by(treatment) legend(order(1 "Observed proportions" 2 "Fitted probabilities"))
> xtitle(Time in months) ytitle(Probability of onycholysis)
```

results in figure 6.8.

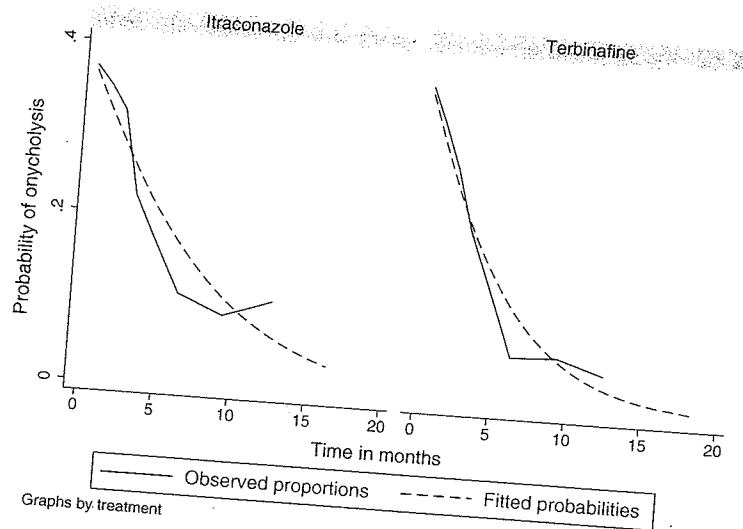


Figure 6.8: Proportions and fitted probabilities using ordinary logistic regression

the margin
However, th
independent
the models
therefore pr
capture bot
dense struc

6.6 Rand

To relax the
patient give
the linear p

with $\zeta_j|x_j$
logistic reg
because it
effect ζ_j .
patient-spe
to onychol
the same v
linear pred

Model
are indepe

Within a t
model as
side) and

Using

where $\zeta_j|$
across bo
patients.

Confi
where ϵ_{ij}
inconsist
(See Skre

egression model
probabilities are

= 1908
= 164.47
= 0.0000
= 0.0830

of. Interval]

7	1.357249
4	.8833679
3	1.006344

mand

(dash)),
probabilities"))

The marginal probabilities predicted by the model fit the observed proportions well. However, the model assumes that the responses for the same patient are conditionally independent given the covariates, which is likely to be violated. As discussed for linear models in section 3.10.1, the standard errors from ordinary logistic regression are therefore probably not trustworthy. A model for longitudinal or clustered data should capture both the *mean structure* (here the marginal probabilities) as well as the *dependence structure*. So far, we have neglected the second aspect but address it now.

6.6 Random-intercept logistic regression

To relax the assumption of conditional independence among the responses for the same patient given the covariates, we can include a patient-specific random intercept ζ_j in the linear predictor

$$\text{logit}\{\Pr(y_{ij}=1|\mathbf{x}_{ij}, \zeta_j)\} = \beta_1 + \beta_2 x_{2j} + \beta_3 x_{3ij} + \beta_4 x_{2j} x_{3ij} + \zeta_j \quad (6.6)$$

with $\zeta_j|\mathbf{x}_{ij} \sim N(0, \psi)$ and ζ_j independent across patients j , giving a random-intercept logistic regression model. This is a simple example of a *generalized linear mixed model* because it is a generalized linear model with both fixed effects β_1 to β_4 and a random effect ζ_j . The random intercept can be thought of as the combined effect of omitted patient-specific (time-constant) covariates that cause some patients to be more prone to onycholysis than others. It is appealing to model this unobserved heterogeneity in the same way as observed heterogeneity by simply adding the random intercept to the linear predictor.

Model specification is completed by assuming that, given $\pi_{ij} \equiv \Pr(y_{ij}|\mathbf{x}_{ij}, \zeta_j)$, y_{ij} are independently distributed as

$$y_{ij}|\pi_{ij} \sim \text{binomial}(1, \pi_{ij})$$

Within a two-stage formulation, Raudenbush and Bryk (2002) refer to this part of the model as the level-1 sampling model and to (6.6) as the structural model (right-hand side) and link function (left-hand side).

Using the latent-response formulation, the model can equivalently be written as

$$y_{ij}^* = \beta_1 + \beta_2 x_{2j} + \beta_3 x_{3ij} + \beta_4 x_{2j} x_{3ij} + \zeta_j + \epsilon_{ij}$$

where $\zeta_j|\mathbf{x}_{ij} \sim N(0, \psi)$ and $\epsilon_{ij}|\mathbf{x}_{ij}, \zeta_j$ has a logistic distribution. The ϵ_{ij} are independent across both occasions and patients and independent of ζ_j , and ζ_j is independent across patients.

Confusingly, logistic random-effects models are sometimes written as $y_{ij} = \pi_{ij} + \epsilon_{ij}$, where ϵ_{ij} is a heteroscedastic normally distributed level-1 residual. This formulation is inconsistent with the random-intercept logistic regression model and should be avoided. (See Skrondal and Rabe-Hesketh 2007.)

6.7 Estimation of logistic random-intercept models

As of Stata 10, there are three commands for fitting the model in Stata: `xtlogit`, `xtmelogit`, and `gllamm`. All three commands provide maximum likelihood estimation using adaptive quadrature to approximate the integrals involved (see sec. 6.11.1 for more information). `xtlogit` follows essentially the same syntax as `xtreg`, `xtmelogit` follows essentially the same syntax as `xtmixed`, and `gllamm` uses essentially the same syntax for linear, logistic, and other types of models.

All three commands are relatively slow because they use numerical integration, but for random-intercept models `xtlogit` is much faster than `xtmelogit`, which is often faster than `gllamm`. However, the rank ordering is reversed when it comes to the usefulness of the commands for predicting random effects and various types of probabilities as we will see in sections 6.12 and 6.13.

We do not discuss random-coefficient logistic regression in this chapter, but such models can be fitted using `xtmelogit` and `gllamm` (but not using `xtlogit`), using essentially the same syntax as for linear random-coefficient models discussed in section 4.5.1. Also, the probit version of the random-intercept model is available using `xtprobit` or `gllamm`, but random-coefficient probit models are available in `gllamm` only.

6.7.1 Using `xtlogit`

The `xtlogit` command for fitting the random-intercept model is similar to the `xtreg` command for fitting the corresponding linear model, except that we add the `quad(30)` option to ensure accurate estimates (see sec. 6.11.1):

```
. xtlogit outcome treatment month trt_month, i(patient) quad(30)
Random-effects logistic regression
Group variable: patient
Random effects u_i ~ Gaussian

Number of obs      =      1908
Number of groups   =       294
Obs per group: min =         1
                  avg =        6.5
                  max =         7

Log likelihood     = -625.38558
Wald chi2(3)       =      150.65
Prob > chi2        =       0.0000
```

outcome	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
treatment	-.160608	.5796716	-0.28	0.782	-1.296744	.9755275
month	-.390956	.0443707	-8.81	0.000	-.4779209	-.3039911
trt_month	-.1367758	.0679947	-2.01	0.044	-.270043	-.0035085
_cons	-1.618795	.4303891	-3.76	0.000	-2.462342	-.7752477
/lnsig2u	2.775749	.1890237			2.405269	3.146228
sigma_u	4.006325	.3786451			3.328876	4.821641
rho	.8298976	.026684			.7710804	.8760322

Likelihood-ratio test of rho=0: $\chi^2(01) = 565.24$ Prob $\geq \chi^2 = 0.000$

6.7.1 Using `xtlogit`

The estimated regression in `sigma_u` represents the and the value next to `rho` latent responses (see sec.

In the itraconazole g decrease by 0.39 per mor 0.14 per month, giving a of time between the two treatment effect, and thi

We can use the or of interpreted as odds ratio the way the results are as "replaying the estima

```
. xtlogit, or
Random-effects logi
Group variable: pat
Random effects u_i
```

Log likelihood =

outcome	
treatment	.
month	.
trt_month	.
/lnsig2u	2
sigma_u	.
rho	.

Likelihood-ratio

The estimated odds ra We see that the estima 0.68 every month and 0.59 (= .6764099 \times .87 odds, 100%(1 - OR), 41% per month in the

odels

in Stata: xtlogit.
likelihood estimation
e sec. 6.11.1 for more
g, xtmelogit follows
ally the same syntax

erical integration. but
logit, which is often
it comes to the useful-
types of probabilities

this chapter, but such
xtlogit), using essen-
cussed in section 4.5.1.
ble using xtprobit or
.lamm only.

is similar to the xtreg
t we add the quad(30)

d(30)
bs = 1908
roups = 294
up: min = 1
avg = 6.5
max = 7
3) = 150.65
2 = 0.0000

[95% Conf. Interval]	
1.296744	.9755275
.4779209	-.3039911
-.270043	-.0035085
2.462342	-.7752477
2.405269	3.146228
3.328876	4.821641
.7710804	.8760322
ob >= chibar2 = 0.000	

The estimated regression coefficients are given in the usual format. The value next to sigma_u represents the estimated standard deviation $\sqrt{\hat{\psi}}$ of the random intercept and the value next to rho represents the estimated residual intraclass correlation of the latent responses (see sec. 6.10.1).

In the itraconazole group (treatment=0), the estimated log odds of onycholysis decrease by 0.39 per month. The log odds for the terbinafine group decrease an extra 0.14 per month, giving a downward slope of 0.53. The estimated difference in the slopes of time between the two groups (the coefficient of trt_month) can be interpreted as the treatment effect, and this is significant at the 5% level.

We can use the or option to obtain exponentiated regression coefficients, which are interpreted as odds ratios here. Instead of refitting the model, we can simply change the way the results are displayed using the following short xtlogit command (known as "replaying the estimation results" in Stata parlance):

```
. xtlogit, or
Random-effects logistic regression
Group variable: patient
Random effects u_i ~ Gaussian

Log likelihood = -625.38558
```

Number of obs	=	1908
Number of groups	=	294
Obs per group: min	=	1
avg	=	6.5
max	=	7
Wald chi2(3)	=	150.65
Prob > chi2	=	0.0000

outcome	OR	Std. Err.	z	P> z	[95% Conf. Interval]	
treatment	.8516258	.4936633	-0.28	0.782	.2734207	2.652566
month	.6764099	.0300128	-8.81	0.000	.6200712	.7378675
trt_month	.8721658	.0593027	-2.01	0.044	.7633467	.9964976
/lnsig2u	2.775749	.1890237			2.405269	3.146228
sigma_u	4.006325	.3786451			3.328876	4.821641
rho	.8298976	.026684			.7710804	.8760322

Likelihood-ratio test of rho=0: chibar2(01) = 565.24 Prob >= chibar2 = 0.000

The estimated odds ratios and their 95% confidence intervals are also given in table 6.2. We see that the estimated odds for a subject in the itraconazole group are multiplied by 0.68 every month and the odds for a subject in the terbinafine group are multiplied by 0.59 (= .6764099 x .872165) every month. In terms of percentage decreases in estimated odds, 100%(1 - OR), the odds decrease 32% per month in the itraconazole group and 41% per month in the terbinafine group.

6.7.2 Using xtmelogit

The syntax for `xtmelogit` is similar to that for `xtmixed` except that we also specify the number of quadrature points, or integration points, using the `intpoints()` option:

```
. xtmelogit outcome treatment month trt_month || patient:, intpoints(30)
Mixed-effects logistic regression
Group variable: patient
```

Number of obs	=	1908
Number of groups	=	294
Obs per group: min	=	1
avg	=	6.5
max	=	7
Integration points	=	30
Log likelihood	=	-625.39709
Wald chi2(3)	=	150.52
Prob > chi2	=	0.0000

outcome	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
treatment	-.1609377	.5842082	-0.28	0.783	-1.305965	.9840893
month	-.3910604	.0443958	-8.81	0.000	-.4780744	-.3040463
trt_month	-.1368073	.0680236	-2.01	0.044	-.270131	-.0034836
_cons	-1.618961	.4347773	-3.72	0.000	-2.471109	-.7668132

Random-effects Parameters		Estimate	Std. Err.	[95% Conf. Interval]	
patient: Identity					
	sd(_cons)	4.008165	.3813919	3.326217	4.829927

LR test vs. logistic regression: $\text{chibar2}(01) = 565.22$ Prob>=chibar2 = 0.0000

The results are similar but not identical to those from `xtlogit` because the commands use slightly different versions of adaptive quadrature (see sec. 6.11.1). Since the estimates took some time to obtain, we store them for later use

```
. estimates store xtmelogit
```

Estimated odds ratios can be obtained using the `or` option. `xtmelogit` can also be used with one integration point, which is equivalent to the so-called Laplace approximation. See section 6.11.2 for the results obtained for the toenail data using this method.

6.7.3 Using gllamm

Using `gllamm` for the random-intercept logistic regression model requires that we specify a logit link and binomial distribution using the `link()` and `family()` options (exactly as for the `glm` command). We also use the `nip()` option (for the number of integration points) to request that 30 integration points be used:

```
. gllamm outcome treatment month trt_month, i(patient) link(logit) family(binomial)
> nlp(30) adapt
number of level 1 units = 1908
number of level 2 units = 294

Condition Number = 23.076299

gllamm model
log likelihood = -625.38558
```

outcome	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
treatment	-.1608751	.5802054	-0.28	0.782	-1.298057	.9763065
month	-.3911055	.0443906	-8.81	0.000	-.4781095	-.3041015
trt_month	-.136829	.0680213	-2.01	0.044	-.2701484	-.0035097
_cons	-1.620364	.4322408	-3.75	0.000	-2.46754	-.7731873

Variances and covariances of random effects

```
***level 2 (patient)
var(1): 16.084107 (3.0626223)
```

The estimates are again similar to those from `xtlogit` and `xtmelogit`. The estimated random-intercept variance is given next to `var(1)` instead of the random-intercept standard deviation reported by `xtlogit` and `xtmelogit`, unless the `variance` option is used for the latter. We store the `gllamm` estimates for later use:

```
. estimates store gllamm
```

We can use the `eform` option to obtain estimated odds ratios or alternatively use the command

```
gllamm, eform
```

after having already fitted the model.

6.8 Inference for logistic random-intercept models

As discussed earlier, we can interpret the regression coefficient β as the difference in log-odds associated with a unit change in the corresponding covariate and the exponentiated regression coefficient as an odds ratio, $OR = \exp(\beta)$. The relevant null hypothesis for odds ratios usually is $H_0: OR = 1$, and this corresponds directly to the null hypothesis that the corresponding regression coefficient is zero, $H_0: \beta = 0$.

Wald tests and z tests can be used for regression coefficients just as described in section 3.6.1 for linear models. 95% Wald confidence intervals for individual regression coefficients are obtained using

Inference

where $z_{.975} = 1$
corresponding c
limits of the co

Wald tests
corresponding
For instance, if
the treatment
20 months is a
sec. 1.8 if this
odds is 0 and 1

```
. lincom
(1) [c
out:
```

If we require
the `lincom` co
limits of the c

```
. lincom
(1) [c
ou'
```

After 20 mc
to itraconaz
so we can s
(= 1/0.055)
Alternativ
of the corre

```
lincom
```

Multivariat
intervals ca
option.

Null hy
ficients can

$$\hat{\beta} \pm z_{.975} \widehat{SE}(\hat{\beta})$$

where $z_{.975} = 1.96$ is the 97.5th percentile of the standard normal distribution. The corresponding confidence interval for the odds ratio is obtained by exponentiating both limits of the confidence interval

$$\exp\{\hat{\beta} - z_{.975} \widehat{SE}(\hat{\beta})\} \text{ to } \exp\{\hat{\beta} + z_{.975} \widehat{SE}(\hat{\beta})\}$$

Wald tests for linear combinations of regression coefficients can be used to test the corresponding multiplicative relationships among odds for different covariate values. For instance, for the toenail data, we may want to obtain the odds ratio comparing the treatment groups after 20 months. The corresponding difference in log odds after 20 months is a linear combination of regression coefficients, namely, $\beta_2 + \beta_4 \times 20$ (see sec. 1.8 if this is not clear). We can test the null hypothesis that the difference in log odds is 0 and hence that the odds ratio is 1 using the `lincom` command:

```
. lincom treatment + trt_month*20
(1) [outcome]treatment + 20 [outcome]trt_month = 0
```

outcome	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
(1)	-2.897456	1.310367	-2.21	0.027	-5.465727	-.3291841

If we require a confidence interval for the odds ratio after 20 months, we can repeat the `lincom` command but this time with the `or` option, which gives exponentials of the limits of the confidence interval above:

```
. lincom treatment + trt_month*20, or
(1) [outcome]treatment + 20 [outcome]trt_month = 0
```

outcome	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
(1)	.0551634	.0722843	-2.21	0.027	.0042293	.7195106

After 20 months of treatment, the odds ratio comparing terbinafine (`treatment=1`) to itraconazole is estimated as 0.055. Such small numbers are difficult to interpret, so we can switch the groups around by taking the reciprocal of the odds ratio, 18 ($= 1/0.055$), which represents the odds ratio comparing itraconazole to terbinafine. Alternatively, we can always switch the comparison around by simply changing the sign of the corresponding difference in log odds in the `lincom` command

```
. lincom -(treatment + trt_month*20), or
```

Multivariate Wald tests can be performed by using `testparm`. Wald tests and confidence intervals can be used with robust standard errors produced by `gllamm` with the `robust` option.

Null hypotheses about individual regression coefficients or several regression coefficients can also be tested using likelihood-ratio tests. Although likelihood-ratio and

Wald tests are asymptotically equivalent, the tests statistics are not identical in finite samples. If the statistics are different, there may be a sparseness problem, for instance with mostly "1" or mostly "0" responses in one of the groups.

Both `xtlogit` and `xtmelogit` provide likelihood-ratio tests for the null hypothesis that the between-cluster variance ψ is zero in the last line of the output. The p -values are based on the correct asymptotic sampling distribution (not the naive χ^2) as described for linear models in section 2.6.2. For the toenail data, the likelihood-ratio statistic is 565.2 giving $p < 0.001$, which suggests that a multilevel model is required.

6.9 Subject-specific vs. population-averaged relationships

The estimated regression coefficients for the random-intercept logistic regression model are more extreme (different from 0) than those for the ordinary logistic regression model (see table 6.2). Correspondingly, the estimated odds ratios are more extreme (different from 1) than those for the ordinary logistic regression model. The reason for this discrepancy is that ordinary logistic regression is fitting overall *population-averaged* probabilities, whereas random-effects logistic regression fits *subject-specific* probabilities for the individual patients.

This important distinction can be seen in the way the two models are written in (6.5) and (6.6). Whereas the former is for the overall or population-averaged probability, conditioning only on covariates, the latter is for the subject-specific probability, given the subject-specific random intercept ζ_j and the covariates. Odds ratios derived from these models can be referred to as population-averaged (although the averaging is applied to the probabilities) or subject-specific odds ratios, respectively. For instance, in the random-intercept model, we can interpret the estimated subject-specific odds ratio of 0.68 for month as the odds ratio for each patient in the itraconazole group: each patient's odds decrease 32% per month. In contrast, the estimated population-averaged odds ratio of 0.84 for month means that the odds of having onycholysis *among the patients* in the itraconazole group, decrease 16% per month. Other commonly used terms for population-averaged and subject-specific are *marginal* and *conditional*, respectively.

The population-averaged probabilities implied by the random-intercept model can be obtained by averaging the subject-specific probabilities over the random-intercept distribution. Since the random intercepts are continuous, this averaging is accomplished by integration:

$$\begin{aligned} \Pr(y_{ij} = 1 | x_{2j}, x_{3ij}) &= \int \Pr(y_{ij} = 1 | x_{2j}, x_{3ij}, \zeta_j) \phi(\zeta_j; 0, \hat{\psi}) d\zeta_j \\ &= \int \frac{\exp(\beta_1 + \beta_2 x_{2j} + \beta_3 x_{3ij} + \beta_4 x_{2j} x_{3ij} + \zeta_j)}{1 + \exp(\beta_1 + \beta_2 x_{2j} + \beta_3 x_{3ij} + \beta_4 x_{2j} x_{3ij} + \zeta_j)} \phi(\zeta_j; 0, \hat{\psi}) d\zeta_j \\ &\neq \frac{\exp(\beta_1 + \beta_2 x_{2j} + \beta_3 x_{3ij} + \beta_4 x_{2j} x_{3ij})}{1 + \exp(\beta_1 + \beta_2 x_{2j} + \beta_3 x_{3ij} + \beta_4 x_{2j} x_{3ij})} \end{aligned} \quad (6.7)$$

where $\phi(\zeta_j; 0, \hat{\psi})$ is the normal density function with mean zero and variance $\hat{\psi}$.

The difference between the average of the average. In the coefficient $\beta_1 + \beta_2 x_{2j} + \beta_3 x_{3ij} + \beta_4 x_{2j} x_{3ij}$ the average of the line is this by comparing average of 1 and 2:

```
. display (invlog
80692783
. display invlog:
81757448
```

We can also see this in specific logistic curve curve represents the :

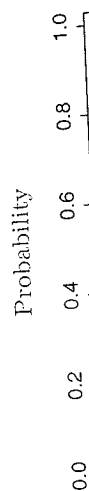


Figure 6.9:

The average effect of x on the curves. However, probability evaluation increasing function

Another way to see the population-averaged model as a latent

The difference between population-averaged and subject-specific effects is due to the fact that the average of a nonlinear function is not the same as the nonlinear function of the average. In the present context, the average of the inverse logit of the linear predictor $\beta_1 + \beta_2 x_{2j} + \beta_3 x_{3ij} + \beta_4 x_{2j} x_{3ij} + \zeta_j$ is not the same as the inverse logit of the average of the linear predictor, which is $\beta_1 + \beta_2 x_{2j} + \beta_3 x_{3ij} + \beta_4 x_{2j} x_{3ij}$. We can see this by comparing the simple average of the logits of 1 and 2 with the logit of the average of 1 and 2:

```
. display (invlogit(1) + invlogit(2))/2
.80592783
. display invlogit((1+2)/1)
.81757448
```

We can also see this in figure 6.9, where the individual dotted curves represent subject-specific logistic curves with randomly varying intercepts, whereas the solid, shallower curve represents the average of these curves for each value of x .

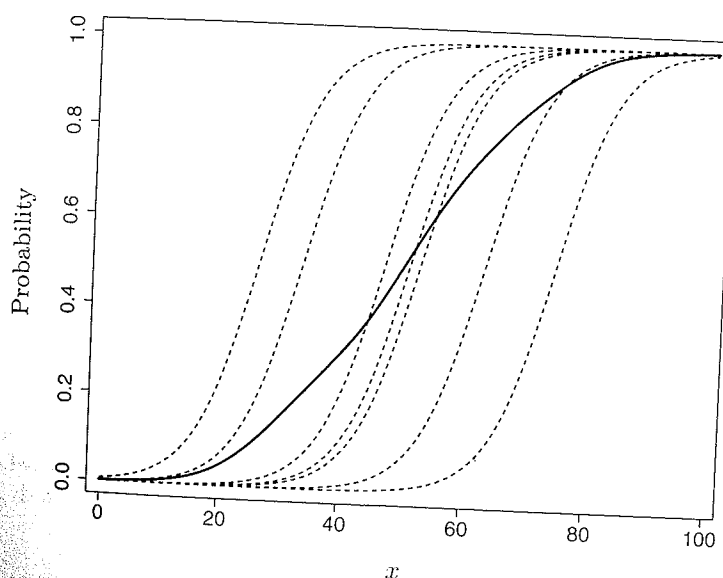


Figure 6.9: Subject-specific versus population-averaged logistic regression

The average curve has a different shape than the individual curves. Specifically, the effect of x on the average curve is smaller than the effect of x on the subject-specific curves. However, the population median probability is the same as the subject-specific probability evaluated at the median of ζ_j because the inverse logit function is a strictly increasing function.

Another way of understanding why the subject-specific effects are more extreme than the population-averaged effects is by writing the random-intercept logistic regression model as a latent-response model: