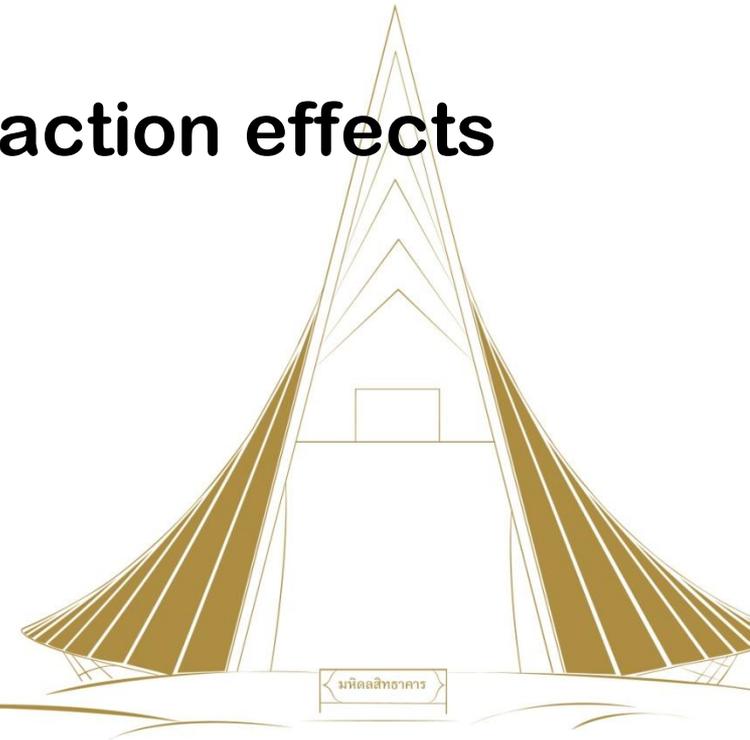




Mahidol University
Wisdom of the Land

Presenting Non-linear interaction effects

Contour plot



Presenter : Narisa Ruenroengbun

Advisor : Ajarn. Anuchate Pattanateepapon

**Department of Clinical Epidemiology and Biostatistics
Faculty of Medicine, Ramathibodi Hospital, Mahidol University**



Outline

- First, I overview the marginal effects framework for summarizing effects in terms of a model's predictions.
- Next, I illustrate the difficulties of testing nonlinear interaction effects even in the context of the linear regression model.
- Then spend some time demonstrating contour plot

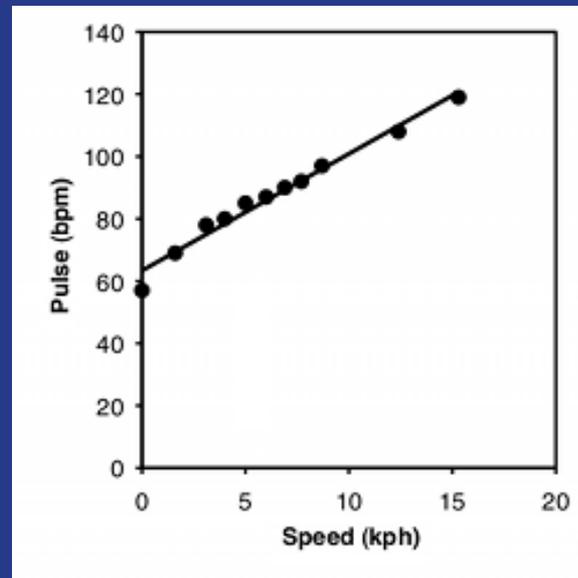


Background

- I. Many effects of interest to sociologists are nonlinear.
- II. Interaction effects exist when the effect of one independent variable is contingent on the level of another independent variable.
- III. For example
 - Does the effect of obtaining more education depend on whether someone is white or black?
 - Does the effect of aging depend on whether someone is a man or a woman?

Linear effects

- Linear effects are those that are constant across the range of a variable



$$Y = a + bX$$

Graph of my pulse rate vs. speed on an elliptical exercise machine.



Linear regression assumption testing

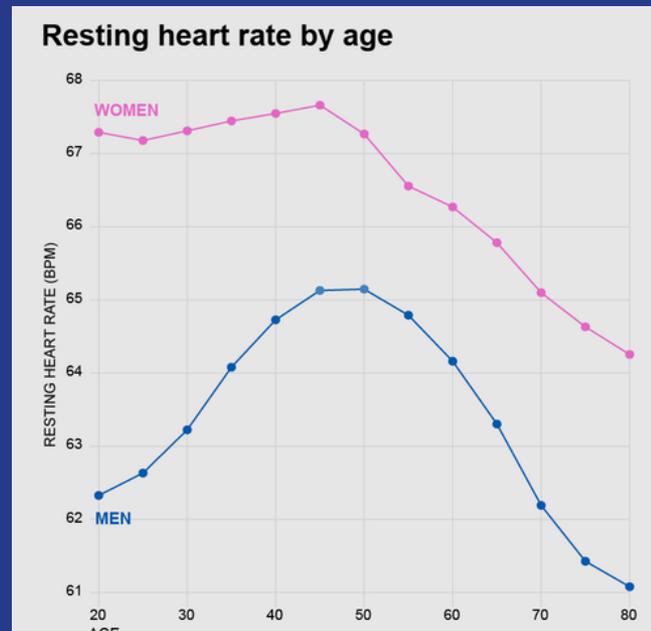
In particular, we will consider the following assumptions.

- **Linearity** – the relationships between the predictors and the outcome variable should be linear
- **Normality** – the errors should be normally distributed
- **Homogeneity of variance (homoscedasticity)** – the error variance should be constant
- **Independence** – the errors associated with one observation are not correlated with the errors of any other observation
- **Errors in variables** – predictor variables are measured without error (we will cover this in Chapter 4)
- **Model specification** – the model should be properly specified (including all relevant variables, and excluding irrelevant variables)



Nonlinear effects

- Nonlinear effects are those that are not constant across the range of an independent variable.
- Example : Age and resting heart rate by age

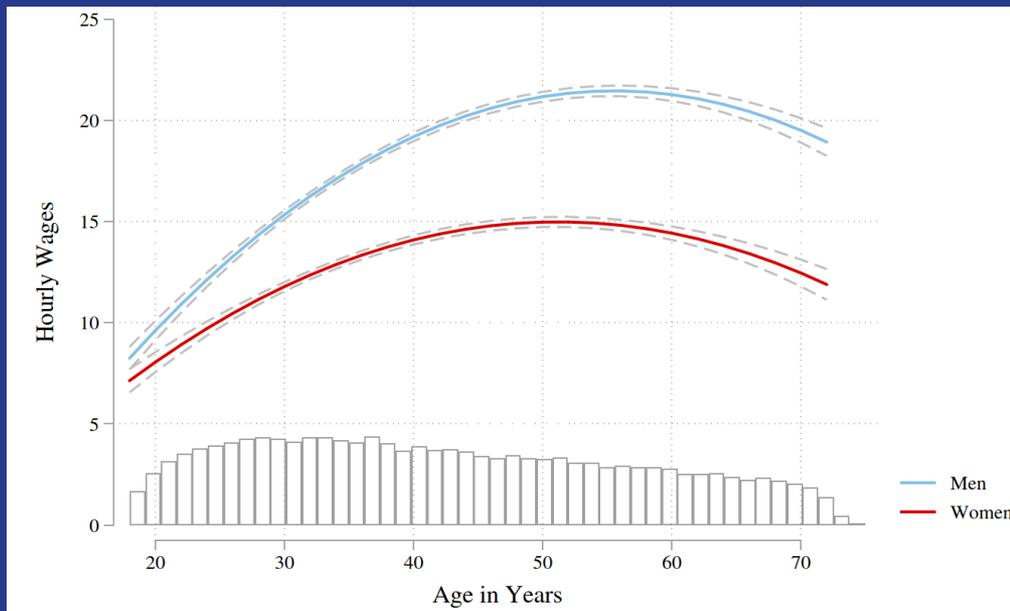




Nonlinear Interaction Effects in the Linear Regression Model

- General Social Survey of employed in UK
- Age likely has a nonlinear relationship with wages.

For example, it is possible that as a person ages his or her wages tend to increase, but this positive effect of age diminishes at older ages and gap of gender



This association suggest that getting older is associated with higher wages, but eventually, the additional wage boost of getting even older begins to diminish.



Detecting non-linearity

- Can do scatterplots
- Can often do incremental F tests or Wald tests
“estat ovtest”

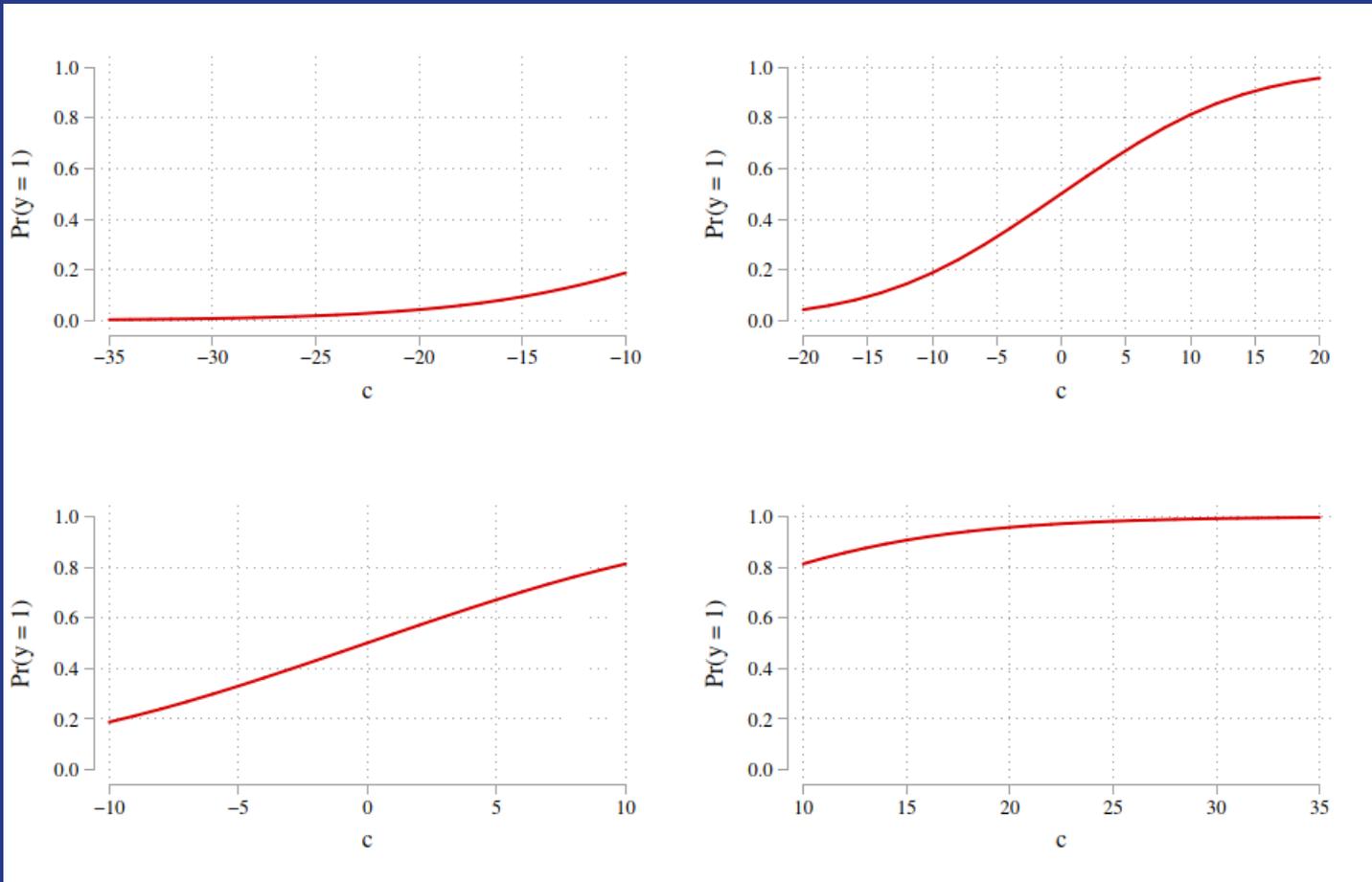


Problem

- There were improper presenting interaction between non-linear relationship
- Primary goal is to provide a set of straightforward best practices for substantive researchers who are interested in presenting interaction effects when the effects of interest are nonlinear.



Simulated data: different effects on the predicted probability of $y = 1$ despite the same coefficient estimate of $\beta = 0.20$.



The regression coefficients do not provide a straightforward summary of the effects in this metric.



Marginal Effects : Summaries of Effects for non-linear relationship

Marginal effects have several advantages over relying on regression coefficients to summarize an independent variable's effect:

1. They allow for one summary measure of an independent variable's effect even when multiple linked coefficients are in the model (e.g., Age and age^2).
2. They avoid the problematic identification (scaling) issues of the coefficients in logit/probit-based
3. On a model's predictions, they can be expressed and interpreted in a different metric than the regression coefficients.



1. Average Marginal Effects (AME)

Average marginal effects are estimated by calculating a marginal effect (ME) for every observation in the sample and then averaging these effects.

First, consider the marginal effect of x_k for a specific observation i :

$$ME_{x_{k_i}} = \eta(x_k = end_i, \mathbf{x} = \mathbf{x}_i) - \eta(x_k = start_i, \mathbf{x} = \mathbf{x}_i).$$

An average marginal effect (AME) is the average (mean) of the marginal effects calculated for each observation in the sample:

$$AME_{x_k} = \frac{1}{N} \sum_{i=1}^N \eta(x_k = end_i, \mathbf{x} = \mathbf{x}_i) - \eta(x_k = start_i, \mathbf{x} = \mathbf{x}_i).$$

An advantage of the AME is each prediction is based on actual observed values in the data.

$x\beta$ is in terms of log odds formula $\eta = \Pr(y=1) = \frac{\exp(\beta x)}{1 + \exp(\beta x)}$

x_k changes from some starting to some ending value



2. Testing the Equality of Marginal Effects: Second Differences

- Testing for interaction effects involves testing the equality of multiple effects. For example,
- When examining possibly interactive effects of the independent variables *age* and *gender*
- **Linear regression** : the coefficient of the product term *ageXgender* provides this test.
- **Nonlinear effects** and/or effects in a different metric than the coefficients, tests of the equality of marginal effects can be used.



2. Testing the Equality of Marginal Effects: Second Differences

$\Delta_{age\ men}$ represent an AME of age for men

$\Delta_{age\ women}$ be the AME of age for women

A test of second difference is a test as to whether two first differences are equal .

A Wald test can be used to determine whether the two effects are equal:

$$z = \frac{\hat{\Delta}_{age\ women} - \hat{\Delta}_{age\ men}}{\sqrt{\hat{\sigma}_{age\ women}^2 + \hat{\sigma}_{age\ men}^2 - 2\hat{\sigma}_{age\ women, age\ men}}}$$

$\hat{\sigma}_{k_b}^2$ is the variance estimate of each marginal effect

$\hat{\sigma}_{k_b}$ is the standard error

$\hat{\sigma}_{k_b=1, k_b=0}$ is the estimate of the covariance between the two effects



MARGIN COMMAND

[R] margins — Marginal means, predictive margins, and marginal effects
([View complete PDF manual entry](#))

Syntax

```
margins [marginlist] [if] [in] [weight] [, response_options options]
```

where *marginlist* is a list of factor variables or interactions that appear in the current estimation results. The variables may be typed with or without the *i.* prefix, and you may use any factor-variable syntax:

```
. margins i.sex i.group i.sex#i.group
. margins sex group sex#i.group
. margins sex##group
```

<i>response_options</i>	Description
-------------------------	-------------

Main

<u>predict</u> (<i>pred_opt</i>)	estimate margins for predict, <i>pred_opt</i>
<u>expression</u> (<i>pnl_exp</i>)	estimate margins for <i>pnl_exp</i>
<u>dydx</u> (<i>varlist</i>)	estimate marginal effect of variables in <i>varlist</i>
<u>eyex</u> (<i>varlist</i>)	estimate elasticities of variables in <i>varlist</i>
<u>dyex</u> (<i>varlist</i>)	estimate semielasticity -- $d(y)/d(\ln x)$
<u>eydx</u> (<i>varlist</i>)	estimate semielasticity -- $d(\ln y)/d(x)$
<u>continuous</u>	treat factor-level indicators as continuous



Result of AME by age and gender

Table 2: Results for how wages are associated with age and gender: tests of average marginal effects (AMEs) and second differences ($N = 30,931$).

	AME_{men}	AME_{women}	Second Difference
<i>Effect of Age</i>			
25 → 30	2.628* (0.081)	1.688* (0.085)	0.940* (0.111)
65 → 70	-1.118* (0.109)	-1.165* (0.120)	0.046 (0.160)
start _{<i>i</i>} + 5	1.149* (0.035)	0.576* (0.035)	0.574* (0.047)

Notes: Standard errors are in parentheses. * $p < 0.05$, two-tailed tests.

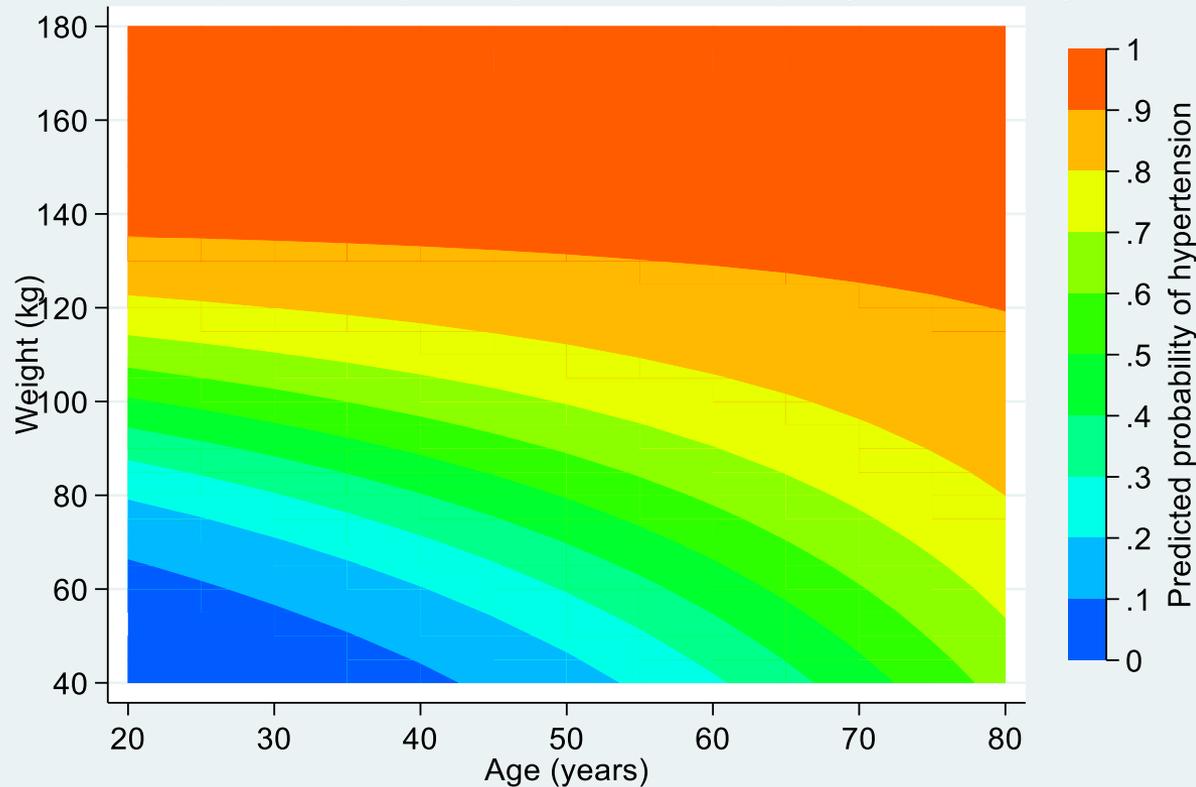


Reporting graphic of non-linear interaction



Contour plot

Predicted probability of hypertension by age and weight





1. Build model by regression

```
. logistic highbp age weight c.age#c.weight
```

```
Logistic regression                Number of obs    =    10,351
LR chi2(3)                        =    2381.23
Prob > chi2                       =    0.0000
Log likelihood = -5860.1512        Pseudo R2       =    0.1689
```

highbp	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
age	1.108531	.0080697	14.15	0.000	1.092827	1.12446
weight	1.081505	.005516	15.36	0.000	1.070748	1.092371
c.age#c.weight	.9992788	.0000977	-7.38	0.000	.9990873	.9994703
_cons	.0002025	.0000787	-21.89	0.000	.0000946	.0004335



2. Estimate probability of outcome with Marginal effect on age

```
. margins, at(age=(20(10)80))

Predictive margins                Number of obs   =   10,351
Model VCE      : OIM

Expression   : Pr(highbp), predict()

1._at       : age           =       20
2._at       : age           =       30
3._at       : age           =       40
4._at       : age           =       50
5._at       : age           =       60
6._at       : age           =       70
7._at       : age           =       80
```

	Delta-method				
	Margin	Std. Err.	z	P> z	[95% Conf. Interval]
_at					
1	.1717358	.0061657	27.85	0.000	.1596512 .1838203
2	.2363955	.0058115	40.68	0.000	.2250051 .2477859
3	.3213383	.0052944	60.69	0.000	.3109615 .3317151
4	.4262457	.0051049	83.50	0.000	.4162404 .4362511
5	.5446655	.0060126	90.59	0.000	.5328809 .55645
6	.6634622	.0075134	88.30	0.000	.6487362 .6781882
7	.7676413	.0082329	93.24	0.000	.751505 .7837775



3. Estimate probability of outcome with Marginal effect on weight

```
. margins, at(weight=(40(20)180))
```

```
Predictive margins          Number of obs   =   10,351
```

```
Model VCE      : OIM
```

```
Expression    : Pr(highbp), predict()
```

```
1._at        : weight      =      40
```

```
2._at        : weight      =      60
```

```
3._at        : weight      =      80
```

```
4._at        : weight      =     100
```

```
5._at        : weight      =     120
```

```
6._at        : weight      =     140
```

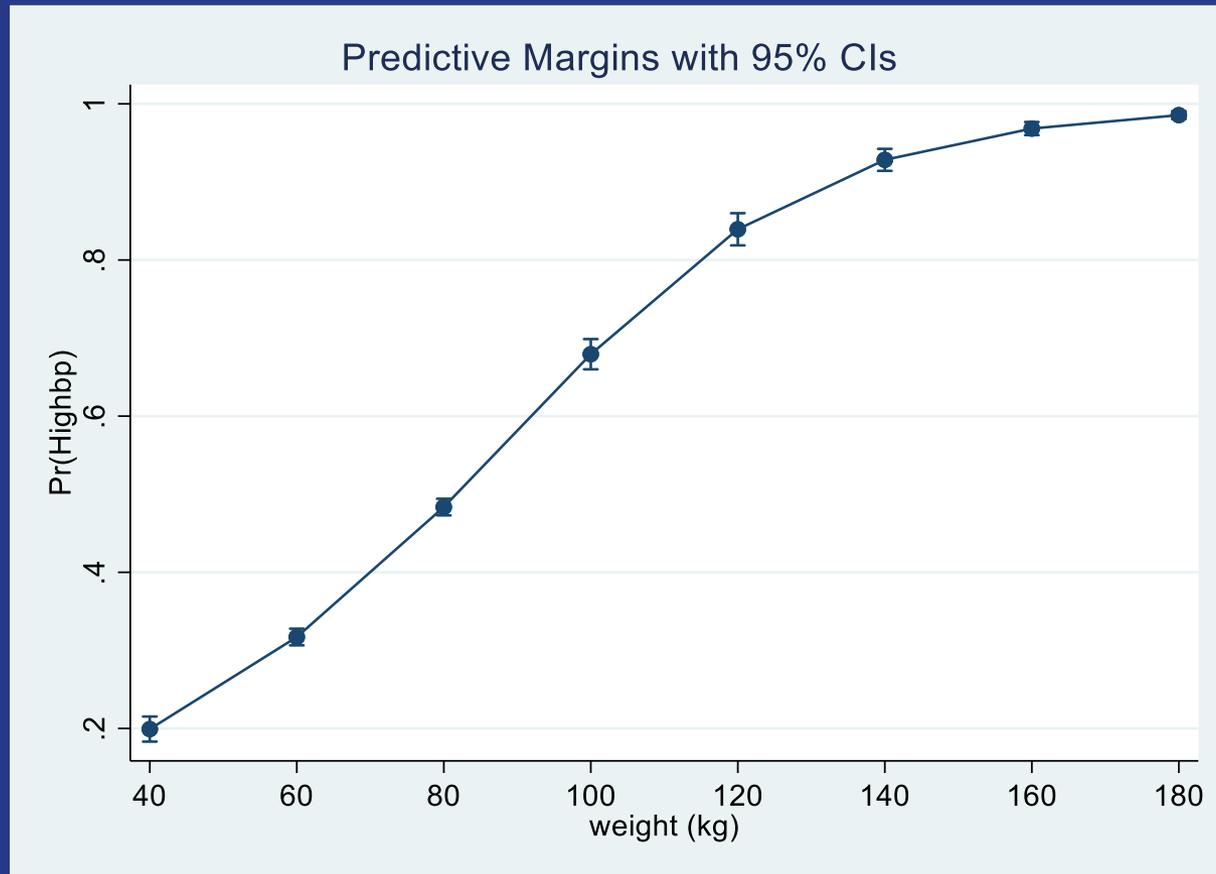
```
7._at        : weight      =     160
```

```
8._at        : weight      =     180
```

	Delta-method				
	Margin	Std. Err.	z	P> z	[95% Conf. Interval]
_at					
1	.1991634	.0082004	24.29	0.000	.183091 .2152358
2	.3170746	.0054472	58.21	0.000	.3063983 .3277508
3	.4835413	.0054383	88.91	0.000	.4728825 .4942002
4	.6793169	.009878	68.77	0.000	.6599564 .6986773
5	.8393498	.0105024	79.92	0.000	.8187656 .8599341
6	.9282807	.0071733	129.41	0.000	.9142213 .9423401
7	.9683396	.0042846	226.01	0.000	.959942 .9767372
8	.9855521	.0025718	383.21	0.000	.9805114 .9905928

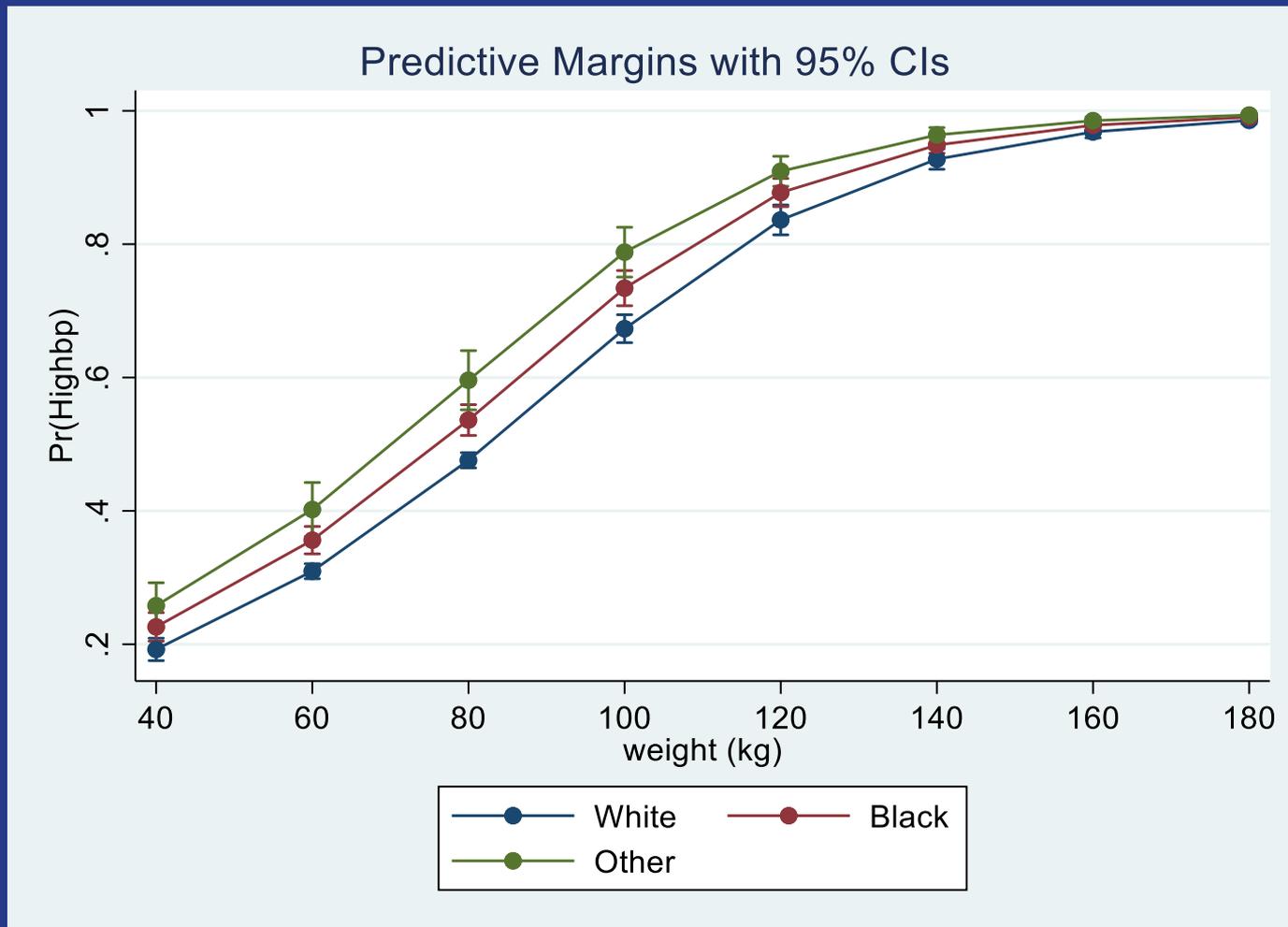


4. Generate predictive plot from marginal model





4. Generate predictive plot from marginal model by interest factor





5. Estimate probability of events of all possible

```
. margins, at(age=(20(5)80) weight=(40(5)180)) saving(predictions, replace)

Adjusted predictions          Number of obs    =    10,351
Model VCE      : OIM

Expression      : Pr(highbp), predict()

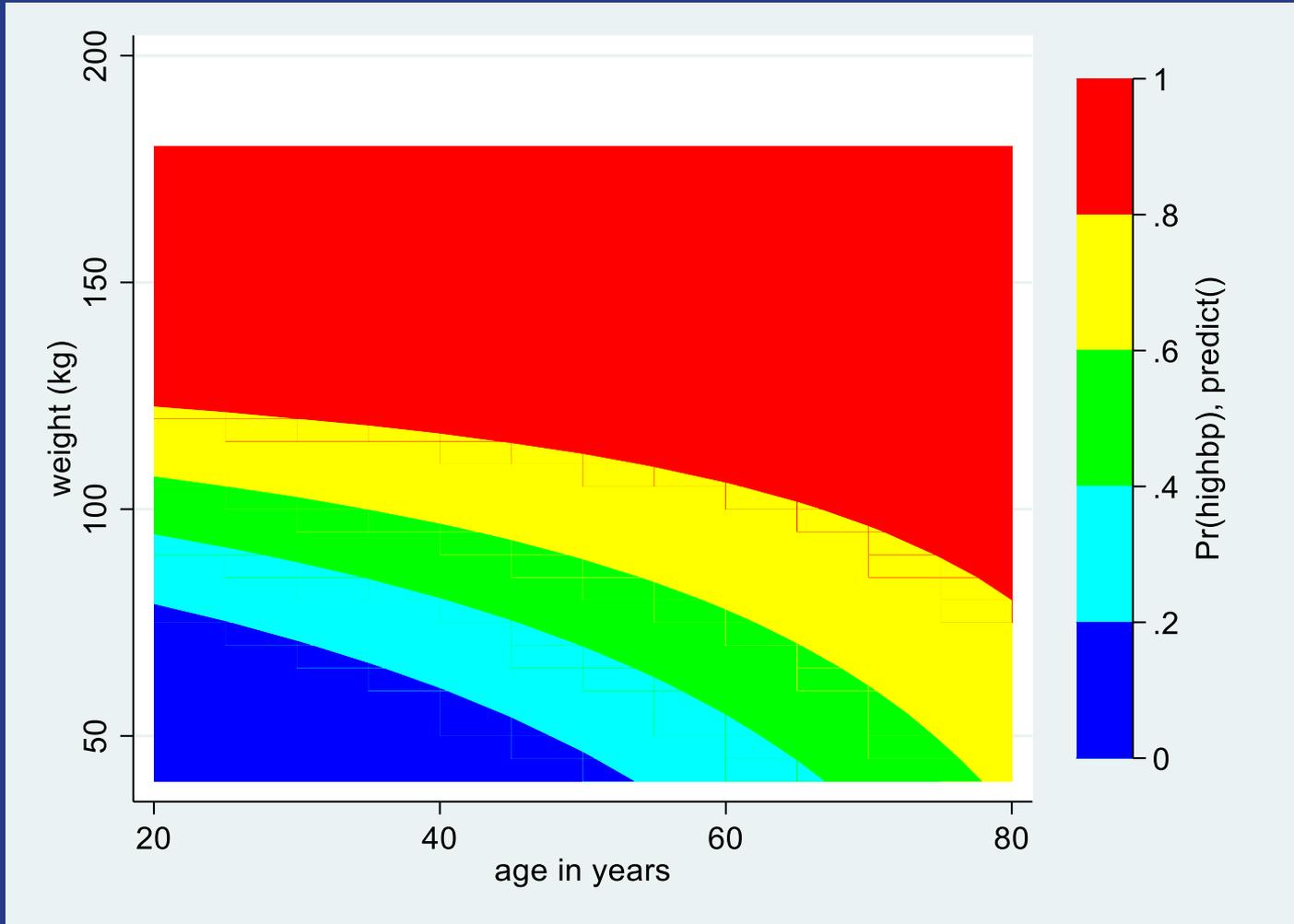
1._at          : age      =      20
                weight    =      40

2._at          : age      =      20
                weight    =      45

3._at          : age      =      20
                weight    =      50
```



6. Construct contour plot present the effect of interaction





Question