Now what do I do with this function?

Enrique Pinzón

StataCorp LP

December 08, 2017
Sao Paulo

Initial thoughts

- Nonparametric regression and about effects/questions
- `npregress`
- Mean relation between an outcome and covariates
  - Model birtweight : age, education level, smoked, number of prenatal visits, ...
  - Model wages: age, education level, profession, tenure, ...
  - $E(y|X)$, conditional mean
- Parametric models have a known functional form

$$
\begin{aligned}
\text{Linear regression:} \quad E(y|X) &= X\beta \\
\text{Binary:} \quad E(y|X) &= F(X\beta) \\
\text{Poisson:} \quad E(y|X) &= \exp(X\beta)
\end{aligned}
$$

- Nonparametric $E(y|X)$. The result of using `predict`

Initial thoughts

- Nonparametric regression and about effects/questions
- `npregress`
- Mean relation between an outcome and covariates
  - Model birtweight : age, education level, smoked, number of prenatal visits, ...
  - Model wages: age, education level, profession, tenure, ...
  - $E(y|X)$, conditional mean
- Parametric models have a known functional form

$$
\begin{aligned}
\text{Linear regression:}\quad E(y|X) &= X\beta \\
\text{Binary:}\quad E(y|X) &= F(X\beta) \\
\text{Poisson:}\quad E(y|X) &= \exp(X\beta)
\end{aligned}
$$

- Nonparametric $E(y|X)$. The result of using `predict`

Initial thoughts

- Nonparametric regression and about effects/questions
- `npregress`
- Mean relation between an outcome and covariates
  - Model birtweight : age, education level, smoked, number of prenatal visits, ...
  - Model wages: age, education level, profession, tenure, ...
  - $E(y|X)$, conditional mean
- Parametric models have a known functional form

$$
\begin{aligned}
\text{Linear regression:} \quad E(y|X) &= X\beta \\
\text{Binary:} \quad E(y|X) &= F(X\beta) \\
\text{Poisson:} \quad E(y|X) &= \exp(X\beta)
\end{aligned}
$$

- Nonparametric $E(y|X)$. The result of using `predict`

Initial thoughts

- Nonparametric regression and about effects/questions
- `npregress`
- Mean relation between an outcome and covariates
  - Model birtweight : age, education level, smoked, number of prenatal visits, ...
  - Model wages: age, education level, profession, tenure, ...
  - $E(y|X)$, conditional mean
- Parametric models have a known functional form

$$\begin{aligned}\text{Linear regression:}\quad E(y|X) &= X\beta\\ \text{Binary:}\quad E(y|X) &= F(X\beta)\\ \text{Poisson:}\quad E(y|X) &= \exp(X\beta)\end{aligned}$$

- Nonparametric $E(y|X)$. The result of using `predict`

Initial thoughts

- Nonparametric regression and about effects/questions
- `npregress`
- Mean relation between an outcome and covariates
    - Model birtweight : age, education level, smoked, number of prenatal visits, ...
    - Model wages: age, education level, profession, tenure, ...
    - $E(y|X)$, conditional mean
- Parametric models have a known functional form

$$
\begin{aligned}
\text{Linear regression:} \quad E(y|X) &= X\beta \\
\text{Binary:} \quad E(y|X) &= F(X\beta) \\
\text{Poisson:} \quad E(y|X) &= \exp(X\beta)
\end{aligned}
$$

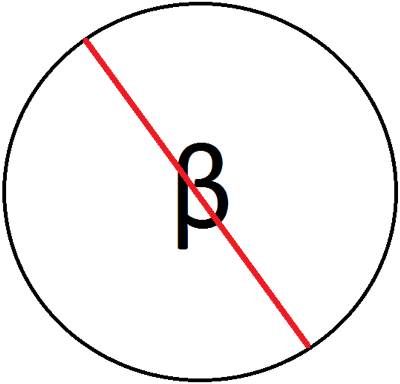
- Nonparametric $E(y|X)$. The result of using `predict`

Initial thoughts

- Nonparametric regression and about effects/questions
- `npregress`
- Mean relation between an outcome and covariates
  - Model birtweight : age, education level, smoked, number of prenatal visits, ...
  - Model wages: age, education level, profession, tenure, ...
  - $E(y|X)$, conditional mean
- Parametric models have a known functional form

$$\begin{aligned} \text{Linear regression:} \quad E(y|X) &= X\beta \\ \text{Binary:} \quad E(y|X) &= F(X\beta) \\ \text{Poisson:} \quad E(y|X) &= \exp(X\beta) \end{aligned}$$

- Nonparametric $E(y|X)$. The result of using `predict`


β

But ...

We had nonparametric regression tools

- lpoly
- lowess

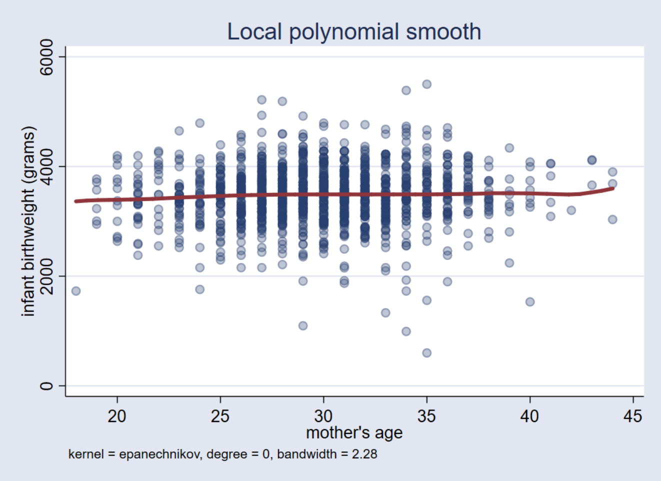
But ...

We had nonparametric regression tools

- `lpoly`
- `lowess`

What happened in the past

```
lpoly bweight mage if (msmoke==0 & medu>12 & fedu>12), ///
      mcolor(%30) lineopts(lwidth(thick))
```



Effects: A thought experiment

I give you the true function

```
. list y x a gx in 1/10, noobs
```

y	x	a	gx
13.46181	.7630615	2	12.73349
1.41086	.9241793	1	1.547555
22.88834	1.816095	2	21.43813
10.97789	.8206556	2	13.01466
11.37173	.0440157	2	10.13213
-.1938587	1.083093	1	.439635
55.87413	3.32037	2	56.56772
2.94979	.8900821	1	1.804343
-1.178733	-2.342678	0	-2.856946
48.79958	3.418333	0	49.94323

Effects: A thought experiment

I give you the true function

```
. list y x a gx in 1/10, noobs
```

y	x	a	gx
13.46181	.7630615	2	12.73349
1.41086	.9241793	1	1.547555
22.88834	1.816095	2	21.43813
10.97789	.8206556	2	13.01466
11.37173	.0440157	2	10.13213
-.1938587	1.083093	1	.439635
55.87413	3.32037	2	56.56772
2.94979	.8900821	1	1.804343
-1.178733	-2.342678	0	-2.856946
48.79958	3.418333	0	49.94323

Effects: A thought experiment

I give you the true function

- Do we know what are the marginal effects
- Do we know causal/treatment effects
- Do we know counterfactuals
- It seems cosmetic
- We cannot use `margins`

Effects: A thought experiment

I give you the true function

- Do we know what are the marginal effects
- Do we know causal/treatment effects
- Do we know counterfactuals
- It seems cosmetic
- We cannot use `margins`

Effects: A thought experiment

I give you the true function

- Do we know what are the marginal effects
- Do we know causal/treatment effects
- Do we know counterfactuals
- It seems cosmetic
- We cannot use `margins`

Effects: A thought experiment

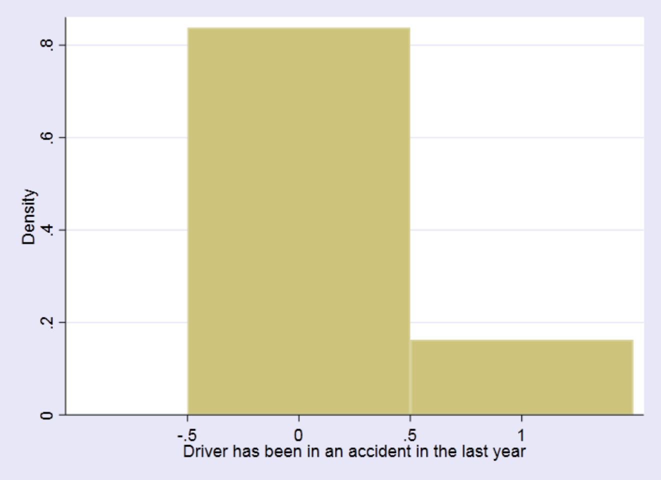
I give you the true function

- Do we know what are the marginal effects
- Do we know causal/treatment effects
- Do we know counterfactuals
- It seems cosmetic
- We cannot use `margins`

A detour

`margins`

Effects: outcome of interest



Data

- **`crash`** 1 if crash
- **`traffic`** Measure of vehicular traffic
- **`tickets`** Number of traffic tickets
- **`male`** 1 if male

Probit model and average marginal effects

probit crash tickets traffic i.male

```
. margins
Predictive margins                              Number of obs     =        948
Model VCE    : OIM
Expression   : Pr(crash), predict()
```

	Delta-method Margin	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_cons	.1626529	.0044459	36.58	0.000	.153939	.1713668

```
. margins, dydx(traffic tickets)
Average marginal effects                        Number of obs     =        948
Model VCE    : OIM
Expression   : Pr(crash), predict()
dy/dx w.r.t. : tickets traffic
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
tickets	.0857818	.0031049	27.63	0.000	.0796963	.0918672
traffic	.0055371	.0020469	2.71	0.007	.0015251	.009549

Probit model and average marginal effects

```
probit crash tickets traffic i.male
```

```
. margins
Predictive margins                              Number of obs     =        948
Model VCE    : OIM
Expression   : Pr(crash), predict()

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   .1626529   .0044459    36.58   0.000      .153939    .1713668
------------------------------------------------------------------------------
```

```
. margins, dydx(traffic tickets)
Average marginal effects                        Number of obs     =        948
Model VCE    : OIM
Expression   : Pr(crash), predict()
dy/dx w.r.t. : tickets traffic

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tickets |   .0857818   .0031049    27.63   0.000     .0796963    .0918672
     traffic |   .0055371   .0020469     2.71   0.007     .0015251     .009549
------------------------------------------------------------------------------
```

Probit model and average marginal effects

```
probit crash tickets traffic i.male
```

```
. margins
Predictive margins                              Number of obs     =        948
Model VCE      : OIM
Expression     : Pr(crash), predict()
```

	Margin	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
_cons	.1626529	.0044459	36.58	0.000	.153939	.1713668

```
. margins, dydx(traffic tickets)
Average marginal effects                        Number of obs     =        948
Model VCE      : OIM
Expression     : Pr(crash), predict()
dy/dx w.r.t. : tickets traffic
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
tickets	.0857818	.0031049	27.63	0.000	.0796963	.0918672
traffic	.0055371	.0020469	2.71	0.007	.0015251	.009549

Not calculus

```
. margins, at(traffic=generate(traffic*1.10)) at(traffic=generate(traffic)) ///
>                 contrast(atcontrast(r) nowald)
Contrasts of predictive margins
Model VCE     : OIM
Expression    : Pr(crash), predict()
1._at         : traffic         = traffic*1.10
2._at         : traffic         = traffic

--------------------------------------------------------------
             |            Delta-method
             |   Contrast   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
         _at |
   (2 vs 1)  |  -.0028589   .0010882     -.0049917   -.0007262
--------------------------------------------------------------
```

Probit model and counterfactuals

```
. margins male
Predictive margins                              Number of obs     =        948
Model VCE    : OIM
Expression   : Pr(crash), predict()
```

	Margin	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
male						
0	.0746963	.0051778	14.43	0.000	.0645481	.0848446
1	.2839021	.008062	35.21	0.000	.2681008	.2997034

```
. margins, dydx(male)
Average marginal effects                        Number of obs     =        948
Model VCE    : OIM
Expression   : Pr(crash), predict()
dy/dx w.r.t. : 1.male
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
1.male	.2092058	.0105149	19.90	0.000	.188597	.2298145

Note: dy/dx for factor levels is the discrete change from the base level.

Probit model and counterfactuals

```
. margins male
Predictive margins                              Number of obs     =        948
Model VCE    : OIM
Expression   : Pr(crash), predict()
------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        male |
          0  |   .0746963   .0051778    14.43   0.000     .0645481    .0848446
          1  |   .2839021    .008062    35.21   0.000     .2681008    .2997034
------------------------------------------------------------------------------

. margins, dydx(male)
Average marginal effects                        Number of obs     =        948
Model VCE    : OIM
Expression   : Pr(crash), predict()
dy/dx w.r.t. : 1.male
------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      1.male |   .2092058   .0105149    19.90   0.000      .188597    .2298145
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.
```

More counterfactuals

```
. margins, dydx(tickets)
Average marginal effects                        Number of obs     =        948
Model VCE    : OIM
Expression   : Pr(crash), predict()
dy/dx w.r.t. : tickets

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tickets |   .0857818   .0031049    27.63   0.000     .0796963    .0918672
------------------------------------------------------------------------------
```

```
. margins, at(tickets=(0(1)5)) contrast(atcontrast(ar) nowald)
Contrasts of predictive margins
Model VCE    : OIM
Expression   : Pr(crash), predict()
1._at        : tickets         =           0
2._at        : tickets         =           1
3._at        : tickets         =           2
4._at        : tickets         =           3
5._at        : tickets         =           4
6._at        : tickets         =           5

--------------------------------------------------------------
             |            Delta-method
             |   Contrast   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
         _at |
   (2 vs 1)  |   .0001208   .0001671     -.0002067    .0004484
   (3 vs 2)  |   .0547975   .0177313      .0200448    .0895502
   (4 vs 3)  |   .3503763   .0225727      .3061346    .3946179
   (5 vs 4)  |    .091227   .0298231      .0327747    .1496793
   (6 vs 5)  |     .37736   .0283876      .3217213    .4329986
--------------------------------------------------------------
```

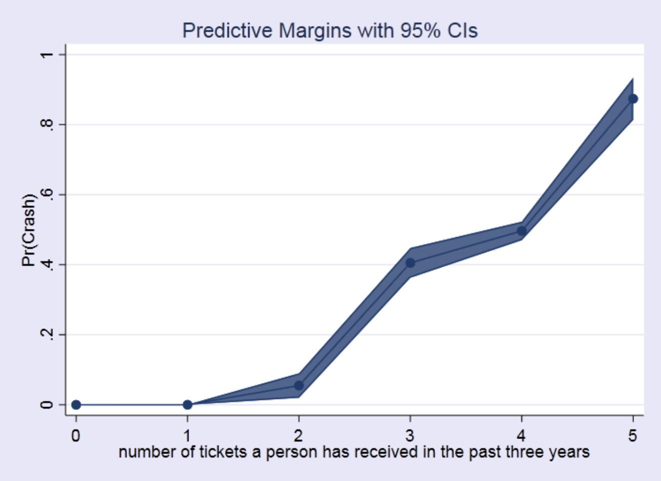
More counterfactuals

```
. margins, dydx(tickets)
Average marginal effects                        Number of obs     =        948
Model VCE     : OIM
Expression    : Pr(crash), predict()
dy/dx w.r.t. : tickets

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     tickets |   .0857818   .0031049    27.63   0.000     .0796963    .0918672
------------------------------------------------------------------------------
```

```
. margins, at(tickets=(0(1)5)) contrast(atcontrast(ar) nowald)
Contrasts of predictive margins
Model VCE     : OIM
Expression    : Pr(crash), predict()
1._at         : tickets         =           0
2._at         : tickets         =           1
3._at         : tickets         =           2
4._at         : tickets         =           3
5._at         : tickets         =           4
6._at         : tickets         =           5

--------------------------------------------------------------
             |            Delta-method
             |   Contrast   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
         _at |
   (2 vs 1)  |   .0001208   .0001671     -.0002067    .0004484
   (3 vs 2)  |   .0547975   .0177313      .0200448    .0895502
   (4 vs 3)  |   .3503763   .0225727      .3061346    .3946179
   (5 vs 4)  |    .091227   .0298231      .0327747    .1496793
   (6 vs 5)  |     .37736   .0283876      .3217213    .4329986
--------------------------------------------------------------
```

marginsplot

margins, at(tickets=(0(1)5))
marginsplot, ciopts(recast(rarea))



Back to nonparametric regression

`npregress` and nonparametric regression

Nonparametric regression: discrete covariates

Mean function for a discrete covariate

- Mean wage conditional on having a college degree

```
. mean wage if collgrad==1
Mean estimation                    Number of obs   =      4,795
```

	Mean	Std. Err.	[95% Conf.	Interval]
wage	8.648064	.0693118	8.512181	8.783947

- `regress wage collgrad, noconstant`
- $E(wage|collgrad = 1)$, nonparametric estimate

Nonparametric regression: discrete covariates

Mean function for a discrete covariate

- Mean wage conditional on having a college degree

```
. mean wage if collgrad==1
Mean estimation                     Number of obs    =      4,795

             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
        wage |   8.648064   .0693118      8.512181    8.783947
```

- `regress wage collgrad, noconstant`
- $E(wage|collgrad = 1)$, nonparametric estimate

Nonparametric regression: discrete covariates

Mean function for a discrete covariate

- Mean wage conditional on having a college degree

```
. mean wage if collgrad==1
Mean estimation                     Number of obs   =      4,795

--------------------------------------------------------------
             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
        wage |   8.648064   .0693118      8.512181    8.783947
--------------------------------------------------------------
```

- `regress wage collgrad, noconstant`
- $E(wage|collgrad = 1)$, nonparametric estimate

Nonparametric regression: continuous covariates

Conditional mean for a continuous covariate

- Mean wage conditional on tenure, measured in years
- $E(wage|tenure = 5.583333)$
- Take observations **near** the value of 5.583333 and then take an average
- $|tenure_i - 5.583333| \leq h$
- h is a small number referred to as the bandwidth

Nonparametric regression: continuous covariates

Conditional mean for a continuous covariate

- Mean wage conditional on tenure, measured in years
- $E(wage|tenure = 5.583333)$
- Take observations **near** the value of 5.583333 and then take an average
- $|tenure_i - 5.583333| \leq h$
- h is a small number referred to as the bandwidth

Nonparametric regression: continuous covariates

Conditional mean for a continuous covariate

- Mean wage conditional on tenure, measured in years
- $E(wage|tenure = 5.583333)$
- Take observations **near** the value of 5.583333 and then take an average
- $|tenure_i - 5.583333| \leq h$
- h is a small number referred to as the bandwidth

Nonparametric regression: continuous covariates

Conditional mean for a continuous covariate

- Mean wage conditional on tenure, measured in years
- $E(wage|tenure = 5.583333)$
- Take observations **near** the value of 5.583333 and then take an average
- $|tenure_i - 5.583333| \leq h$
- h is a small number referred to as the bandwidth

Nonparametric regression: continuous covariates

Conditional mean for a continuous covariate

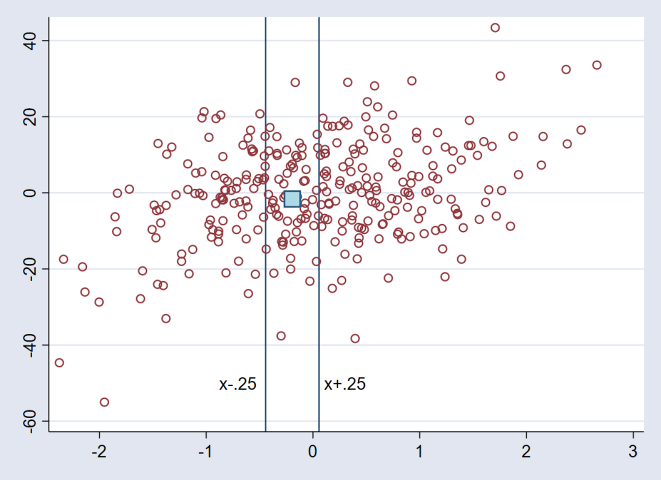
- Mean wage conditional on tenure, measured in years
- $E(wage|tenure = 5.583333)$
- Take observations **near** the value of 5.583333 and then take an average
- $|tenure_i - 5.583333| \leq h$
- h is a small number referred to as the bandwidth

Nonparametric regression: continuous covariates

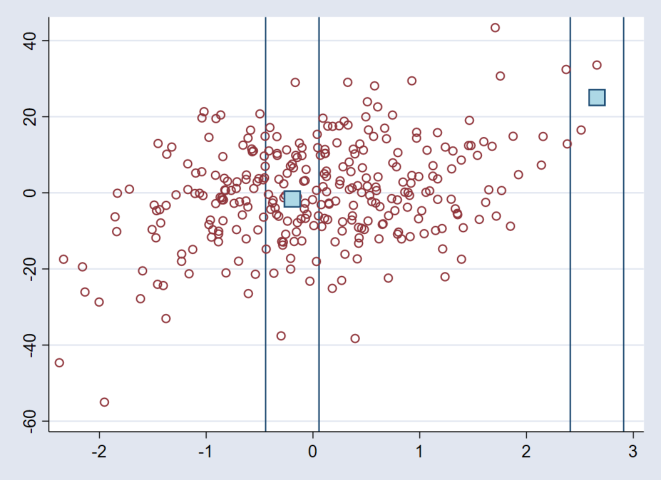
Conditional mean for a continuous covariate

- Mean wage conditional on tenure, measured in years
- $E(wage|tenure = 5.583333)$
- Take observations **near** the value of 5.583333 and then take an average
- $|tenure_i - 5.583333| \leq h$
- h is a small number referred to as the bandwidth

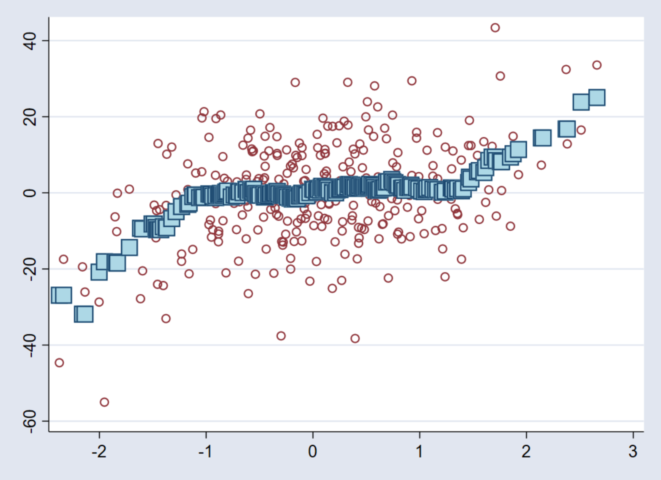
Graphical example



Graphical example



Graphical example continued



Two concepts

1. h
2. Definition of distance between points, $\left|\frac{x_i - x}{h}\right| \leq 1$

Kernel weights

$u \equiv \frac{x_i - x}{h}$

Kernel	$K(u)$
Gaussian	$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right)$
Epanechnikov	$\frac{3}{4\sqrt{5}}\left(1-\frac{u^2}{5}\right)\mathbb{I}\left(\lvert u\rvert \leq \sqrt{5}\right)$
Epanechnikov2	$\frac{3}{4}\left(1-u^2\right)\mathbb{I}\left(\lvert u\rvert \leq 1\right)$
Rectangular(Uniform)	$\frac{1}{2}\mathbb{I}\left(\lvert u\rvert \leq 1\right)$
Triangular	$\left(1-\lvert u\rvert\right)\mathbb{I}\left(\lvert u\rvert \leq 1\right)$
Biweight	$\frac{15}{16}\left(1-u^2\right)^2\mathbb{I}\left(\lvert u\rvert \leq 1\right)$
Triweight	$\frac{35}{32}\left(1-u^2\right)^3\mathbb{I}\left(\lvert u\rvert \leq 1\right)$
Cosine	$\left(1+\cos\left(2\pi u\right)\right)\mathbb{I}\left(\lvert u\rvert \leq \frac{1}{2}\right)$
Parzen	$\left(\frac{4}{3}-8u^2+8\lvert u\rvert^3\right)\mathbb{I}\left(\lvert u\rvert \leq \frac{1}{2}\right) + \frac{8}{3}\left(1-\lvert u\rvert\right)^3\mathbb{I}\left(\frac{1}{2} < \lvert u\rvert \leq 1\right)$

Kernel weights

$u \equiv \frac{x_i - x}{h}$

Kernel	$K(u)$
Gaussian	$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right)$
Epanechnikov	$\frac{3}{4\sqrt{5}}\left(1-\frac{u^2}{5}\right)\mathbb{I}\left(\lvert u\rvert \leq \sqrt{5}\right)$
Epanechnikov2	$\frac{3}{4}\left(1-u^2\right)\mathbb{I}\left(\lvert u\rvert \leq 1\right)$
Rectangular(Uniform)	$\frac{1}{2}\mathbb{I}\left(\lvert u\rvert \leq 1\right)$
Triangular	$\left(1-\lvert u\rvert\right)\mathbb{I}\left(\lvert u\rvert \leq 1\right)$
Biweight	$\frac{15}{16}\left(1-u^2\right)^2\mathbb{I}\left(\lvert u\rvert \leq 1\right)$
Triweight	$\frac{35}{32}\left(1-u^2\right)^3\mathbb{I}\left(\lvert u\rvert \leq 1\right)$
Cosine	$\left(1+\cos\left(2\pi u\right)\right)\mathbb{I}\left(\lvert u\rvert \leq \frac{1}{2}\right)$
Parzen	$\left(\frac{4}{3}-8u^2+8\lvert u\rvert^3\right)\mathbb{I}\left(\lvert u\rvert \leq \frac{1}{2}\right)$ $+\frac{8}{3}\left(1-\lvert u\rvert\right)^3\mathbb{I}\left(\frac{1}{2} < \lvert u\rvert \leq 1\right)$

Discrete bandwidths

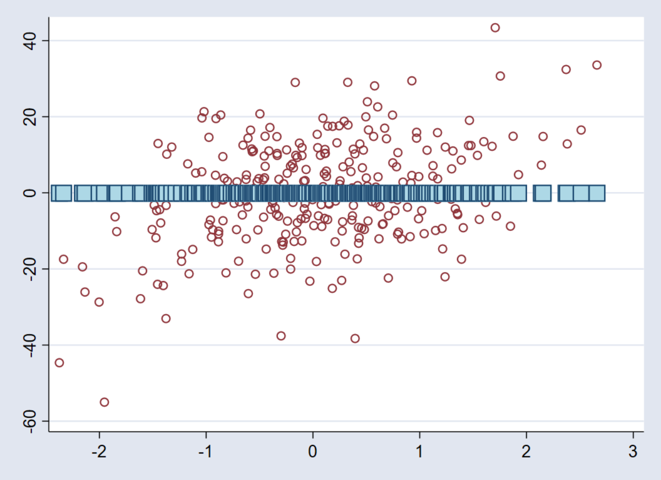
- Default

$$k\left(.\right) = \begin{cases} 1 & \text{if} \quad x_i = x \\ h & \text{otherwise} \end{cases}$$

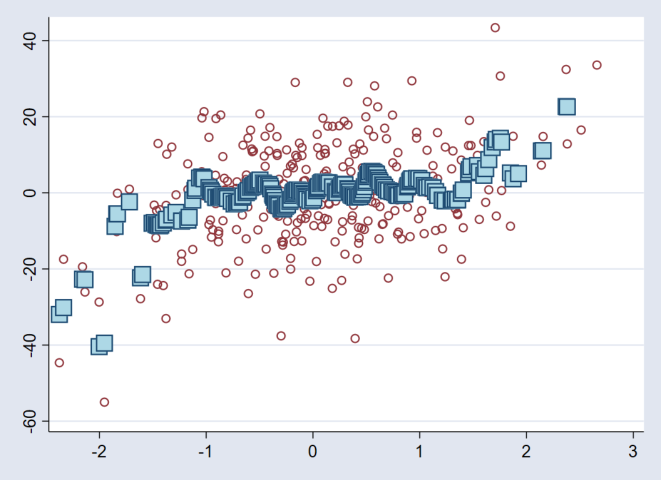
- Cell mean

$$k\left(.\right) = \begin{cases} 1 & \text{if} \quad x_i = x \\ 0 & \text{otherwise} \end{cases}$$

Bandwidth (bias)



Bandwidth (variance)



Estimation

- Choose bandwidth optimally. Minimize bias–variance trade–off
  - Cross-validation (default)
  - Improved AIC (IMAIC)
- Compute a regression for every point in data (local linear)
  - Computes derivatives and derivative bandwidths
- Compute a mean for every point in data (local-constant)

Example

- `citations` monthly drunk driving citations
- `taxes` 1 if alcoholic beverages are taxed
- `fines` drunk driving fines in thousands
- `csize` city size (small, medium, large)
- `college` 1 if college town

Example

- `citations` monthly drunk driving citations
- `taxes` 1 if alcoholic beverages are taxed
- `fines` drunk driving fines in thousands
- `csize` city size (small, medium, large)
- `college` 1 if college town

`npregress` bandwidth

```
. npregress kernel citations fines

Computing mean function

Minimizing cross-validation function:

Iteration 0:   Cross-validation criterion =  35.478784
Iteration 1:   Cross-validation criterion =  4.0147129
Iteration 2:   Cross-validation criterion =  4.0104176
Iteration 3:   Cross-validation criterion =  4.0104176
Iteration 4:   Cross-validation criterion =  4.0104176
Iteration 5:   Cross-validation criterion =  4.0104176
Iteration 6:   Cross-validation criterion =  4.0104006

Computing optimal derivative bandwidth

Iteration 0:   Cross-validation criterion =  6.1648059
Iteration 1:   Cross-validation criterion =  4.3597488
Iteration 2:   Cross-validation criterion =  4.3597488
Iteration 3:   Cross-validation criterion =  4.3597488
Iteration 4:   Cross-validation criterion =  4.3597488
Iteration 5:   Cross-validation criterion =  4.3597488
Iteration 6:   Cross-validation criterion =  4.3595842
Iteration 7:   Cross-validation criterion =  4.3594713
Iteration 8:   Cross-validation criterion =  4.3594713
```

npregress output

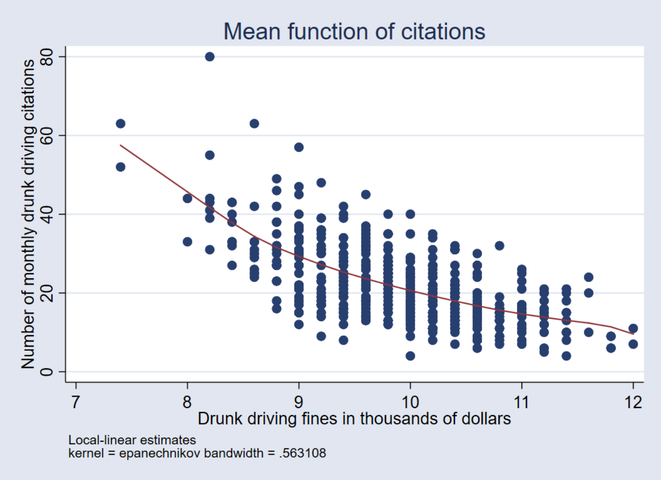
```
. npregress kernel citations fines, nolog
Bandwidth
──────────────────────────────────────
             │       Mean      Effect
─────────────┼────────────────────────
Mean         │
       fines │   .5631079     .924924
──────────────────────────────────────

Local-linear regression                    Number of obs      =        500
Kernel   : epanechnikov                    E(Kernel obs)      =        282
Bandwidth: cross validation                R-squared          =     0.4380
──────────────────────────────────────────────────────────────────────────
   citations │   Estimate
─────────────┼────────────────────────────────────────────────────────────
Mean         │
   citations │   22.33999
─────────────┼────────────────────────────────────────────────────────────
Effect       │
       fines │  -7.692388
──────────────────────────────────────────────────────────────────────────
Note: Effect estimates are averages of derivatives.
Note: You may compute standard errors using vce(bootstrap) or reps().
```

npregress predicted values

```
. describe _*
               storage   display    value
variable name    type    format     label      variable label
--------------------------------------------------------------------------------
_Mean_citations double   %10.0g                mean function
_d_Mean_citat~s double   %10.0g                derivative of mean function w.r.t
                                                 fines
```

npgraph

```
. npgraph
```



`npregress` standard errors I

```
 . quietly npregress kernel citations fines, reps(3) seed(111)
. estimates store A
. quietly npregress kernel citations fines, vce(bootstrap, reps(3) seed(111))
. estimates store B
. estimates table A B, se

----------------------------------------------
    Variable |        A             B
-------------+--------------------------------
Mean         |
   citations |   22.339995     22.339995
             |   .65062763     .65062763
-------------+--------------------------------
Effect       |
       fines |  -7.6923878    -7.6923878
             |   .23195785     .23195785
----------------------------------------------
                                 legend: b/se
```

`npregress` standard errors II (percentile C.I.)

```
. npregress
Bandwidth
─────────────────────────────────────
              │      Mean     Effect
──────────────┼──────────────────────
Mean          │
        fines │  .5631079    .924924
─────────────────────────────────────

Local-linear regression                    Number of obs      =        500
Kernel   : epanechnikov                    E(Kernel obs)      =        282
Bandwidth: cross validation                R-squared          =     0.4380
──────────────────────────────────────────────────────────────────────────
              │  Observed   Bootstrap                        Percentile
    citations │  Estimate   Std. Err.      z    P>|z|     [95% Conf. Interval]
──────────────┼───────────────────────────────────────────────────────────
Mean          │
    citations │  22.33999   .6506276    34.34   0.000     21.54051    22.74807
──────────────┼───────────────────────────────────────────────────────────
Effect        │
        fines │ -7.692388   .2319578   -33.16   0.000    -7.701931   -7.267385
──────────────────────────────────────────────────────────────────────────
Note: Effect estimates are averages of derivatives.
```

A more interesting model

```
 . npregress kernel citations fines i.taxes i.csize i.college,
              reps(200) seed(10)
Bandwidth
───────────────────────────────────────────
             │      Mean     Effect
─────────────┼─────────────────────────────
Mean         │
       fines │   .4471373   .6537197
       taxes │   .4375656   .4375656
       csize │   .3938759   .3938759
     college │    .554583    .554583
───────────────────────────────────────────
Local-linear regression                    Number of obs      =         500
Continuous kernel : epanechnikov           E(Kernel obs)      =         224
Discrete kernel   : liracine               R-squared          =      0.8010
Bandwidth         : cross validation
─────────────────────────────────────────────────────────────────────────────
             │   Observed   Bootstrap                         Percentile
   citations │   Estimate   Std. Err.      z    P>|z|     [95% Conf. Interval]
─────────────┼───────────────────────────────────────────────────────────────
Mean         │
   citations │   22.26306   .4616724    48.22   0.000     21.39581    23.30278
─────────────┼───────────────────────────────────────────────────────────────
Effect       │
       fines │  -7.332833   .3341222   -21.95   0.000    -7.970275   -6.665263
             │
       taxes │
        (tax │
          vs │
    no tax)  │  -4.502718   .4946306    -9.10   0.000    -5.360078   -3.465397
             │
       csize │
     (medium │
          vs │
      small) │   5.300524   .2731374    19.41   0.000     4.723821    5.879301
      (large │
          vs │
      small) │   11.05053   .5236424    21.10   0.000     9.942253     12.1252
             │
     college │
    (college │
          vs │
  not coll..)│   5.953188    .500154    11.90   0.000     4.937102    6.969837
─────────────────────────────────────────────────────────────────────────────
Note: Effect estimates are averages of derivatives for continuous covariates
      and averages of contrasts for factor covariates.
```

margins

```
. margins, at(fines=generate(fines)) at(fines=generate(fines*1.15)) ///
>         contrast(atcontrast(r) nowald) reps(200) seed(12)
(running margins on estimation sample)
Bootstrap replications (200)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
..................................................   100
..................................................   150
..................................................   200
Contrasts of predictive margins
                                              Number of obs     =        500
                                              Replications      =        200

Expression   : mean function, predict()
1._at        : fines          = fines
2._at        : fines          = fines*1.15

-------------------------------------------------------------
             |   Observed   Bootstrap         Percentile
             |   Contrast   Std. Err.     [95% Conf. Interval]
-------------+-----------------------------------------------
         _at |
   (2 vs 1)  |  -8.254875   .8058215     -10.44121   -7.381583
-------------------------------------------------------------
```

Another example with `margins`

$$y = \begin{cases} 10 + x^3 + \varepsilon & \text{if} \quad a = 0 \\ 10 + x^3 - 10x + \varepsilon & \text{if} \quad a = 1 \\ 10 + x^3 + 3x + \varepsilon & \text{if} \quad a = 2 \end{cases}$$

Mean and marginal effects

```
. quietly regress y (c.x#c.x#c.x)#i.a c.x#i.a
. margins
Predictive margins                              Number of obs     =      1,000
Model VCE    : OLS
Expression   : Linear prediction, predict()

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   12.02269   .0313857   383.06   0.000      11.9611    12.08428
------------------------------------------------------------------------------
```

```
. margins, dydx(*)
Average marginal effects                        Number of obs     =      1,000
Model VCE    : OLS
Expression   : Linear prediction, predict()
dy/dx w.r.t. : 1.a 2.a x

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           a |
          1  |  -9.781302     .05743  -170.32   0.000       -9.894   -9.668604
          2  |   3.028531   .0544189    55.65   0.000     2.921742     3.13532
             |
           x |    3.97815   .0303517   131.07   0.000      3.91859    4.037711
------------------------------------------------------------------------------
Note: dy/dx for factor levels is the discrete change from the base level.
```

Mean and marginal effects

```
. quietly regress y (c.x#c.x#c.x)#i.a c.x#i.a
. margins
Predictive margins                             Number of obs     =      1,000
Model VCE    : OLS
Expression   : Linear prediction, predict()
```

	Margin	Delta-method Std. Err.	t	P>|t|	[95% Conf.	Interval]
_cons	12.02269	.0313857	383.06	0.000	11.9611	12.08428

```
. margins, dydx(*)
Average marginal effects                       Number of obs     =      1,000
Model VCE    : OLS
Expression   : Linear prediction, predict()
dy/dx w.r.t. : 1.a 2.a x
```

	dy/dx	Delta-method Std. Err.	t	P>|t|	[95% Conf.	Interval]
a						
1	-9.781302	.05743	-170.32	0.000	-9.894	-9.668604
2	3.028531	.0544189	55.65	0.000	2.921742	3.13532
x	3.97815	.0303517	131.07	0.000	3.91859	4.037711

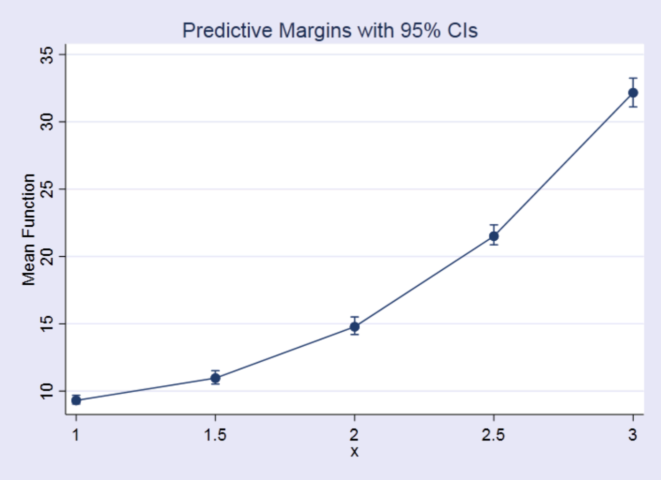
Note: dy/dx for factor levels is the discrete change from the base level.

`npregress` estimates

```
. npregress kernel y x i.a, vce(bootstrap, reps(100) seed(111))
(running npregress on estimation sample)
Bootstrap replications (100)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
..................................................   100
Bandwidth

------------------------------------
             |      Mean     Effect
-------------+----------------------
Mean         |
           x |  .3630656   .5455175
           a |  3.05e-06   3.05e-06
------------------------------------

Local-linear regression                 Number of obs      =          1,000
Continuous kernel : epanechnikov        E(Kernel obs)      =            363
Discrete kernel   : liracine            R-squared          =         0.9888
Bandwidth         : cross validation
------------------------------------------------------------------------------
             |   Observed   Bootstrap                          Percentile
           y |   Estimate   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Mean         |
           y |   12.34335   .3195918    38.62   0.000     11.57571    12.98202
-------------+----------------------------------------------------------------
Effect       |
           x |   3.619627   .2937529    12.32   0.000     3.063269    4.143166
             |
           a |
   (1 vs 0)  |  -9.881542   .3491042   -28.31   0.000     -10.5277   -9.110781
   (2 vs 0)  |   3.168084   .2129506    14.88   0.000      2.73885    3.570004
------------------------------------------------------------------------------
Note: Effect estimates are averages of derivatives for continuous covariates
      and averages of contrasts for factor covariates.
```

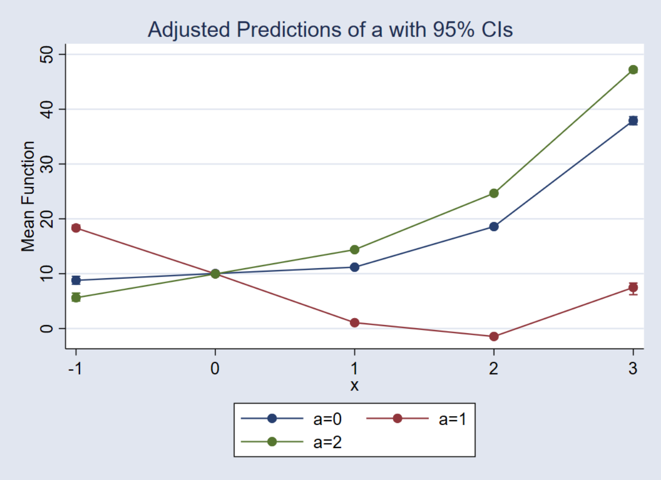
Function for different values of `x`

```
. margins, at(x=(1(.5)3)) reps(100) seed(111)
```



Funtion at different values of `x` for all `a`

```
. margins a, at(x=(-1(1)3)) reps(100) seed(111)
```



Conclusion

- Intuition about nonparametric regression
- Details about how `npregress`
- Importance of being able to ask questions to your model