

# Regresiones por cuantiles para un análisis de duración: Desempleo en el Perú

Roger Asencios

# Motivación

## Duración de desempleo 2002 -2013 (Semanas promedio)

	Duración Incompleta
<b>Sexo</b>	
Hombre	11.8
Mujer	11.3
<b>Edad</b>	
14 - 24	10.6
25 - 44	12.0
45 a más	12.8
<b>Nivel educativo</b>	
Inicial o sin nivel	11.2
Secundia y primaria	11.1
Educacion superior	11.9
<b>Total</b>	<b>11.5</b>

Fuente: Encuesta Permanente Empleo

## Motivación

- Generalmente el efecto de estas variables en la duración del desempleo, se ha modelado con el uso de modelos de riesgo proporcional

$$\lambda(t; \mathbf{x}) = k(\mathbf{x})\lambda_0(t)$$

- Produce un resumen general de los promedios de la distribución correspondiente a un conjunto de  $\mathbf{x}$ 's. Nosotros podríamos ir más allá y calcular varias curvas de regresión que correspondan a distintos puntos de la distribución y obtener una imagen más completa del conjunto.

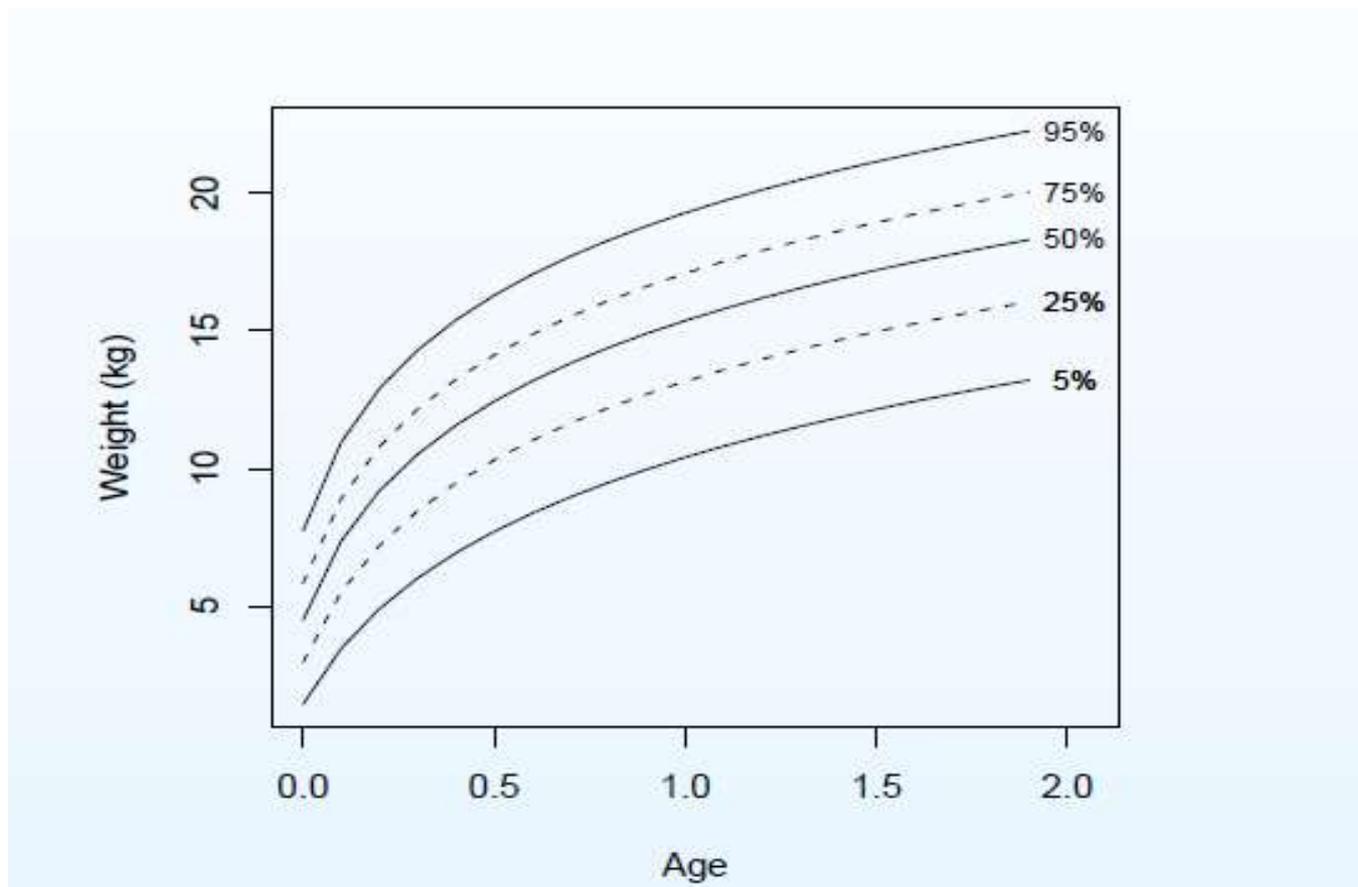
## Motivación

- Usar una metodología que permita mostrar efectos que, de ser usado el modelo convencional, permanecerían ocultos.

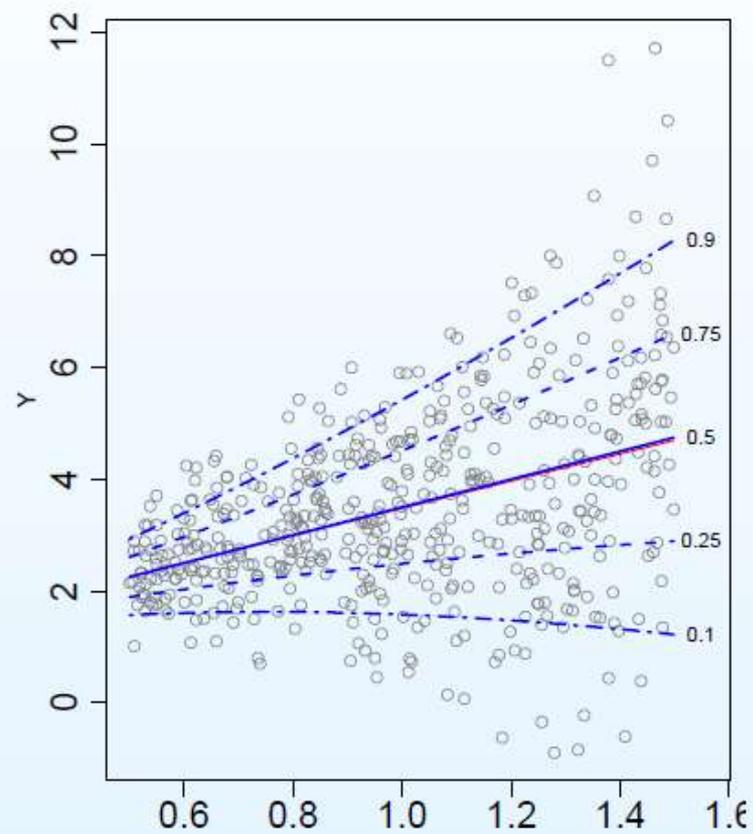
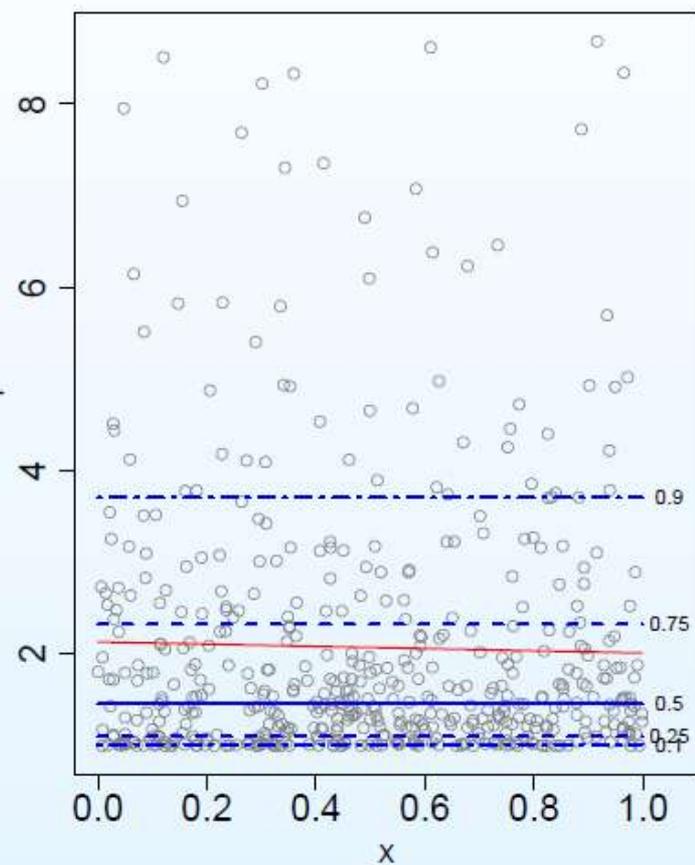
$$Q_{\log(Y_i)/x_i}(\tau/x_i) = x_i' \beta(\tau)$$

- Permite que los efectos de las variables explicativas varíen en los diferentes puntos de la distribución. Con esto se puede capturar, por ejemplo variables que afectan de manera importante en una cola de la distribución, pero no en la otra.

## Regresiones por cuantiles



# Regresiones por cuantiles



## Regresiones por cuantiles

- Los cuantiles están inseparablemente ligados a las operaciones de sortear y ordenar las observaciones de una muestra. Pero se puede definir los cuantiles a través de un problema de optimización.
- Koenker y Bassett (1978) definen el  $\tau$ th cuantile muestral como una solución a:

$$\hat{\alpha}(\tau) = \underset{a \in \mathbb{R}}{\operatorname{argmin}} \sum_{i=1}^n \rho_{\tau}(y_i - a)$$

## Regresiones por cuantiles

- Donde  $\rho_{\tau}(\mathbf{u}) = (\tau - I(\mathbf{u} \leq \mathbf{0}))$  and  $I(\cdot)$  es una función indicador.

- Cuando  $\tau = \frac{1}{2}$

$$\text{mediana} = \underset{a \in R}{\operatorname{argmin}} \sum_{i=1}^n |y_i - a|$$

- El modelo linear para el  $\tau$ th quantile

$$y_i = x_i^T \beta_{\tau} + e_i, \quad i = 1, \dots, n,$$

## Regresiones por cuantiles

- Donde el  $\tau$ th cuantile de  $e_i$  es cero.

$$Q_{\tau}(Y/X) = X^T \beta_{\tau}$$

- Y el estimador de  $\beta_{\tau}$

$$\widehat{\beta}_{\tau} = \operatorname{argmin}_{\beta \in R^p} \sum_{i=1}^n \rho_{\tau}(y_i - x_i^T \beta)$$

## Revision de literatura

- Chacaltana (1999) encuentra que a más tiempo desempleado, la probabilidad de dejar el desempleo es mayor. Variables como ser hombre, ser cesante o tener estudios primarios afectan positivamente el riesgo de salir del desempleo. Variables como la edad, jefe del hogar y tener estudios superiores afectan negativamente esta probabilidad.
- Díaz y Maruyama (2000) encuentran que la probabilidad de salir del desempleo aumenta con la duración del episodio, y que variables como experiencia laboral previa y tener educación primaria o secundaria aumentan la probabilidad de salir del desempleo, mientras que a mayor edad y mayor ingreso familiar la probabilidad de salida es menor.

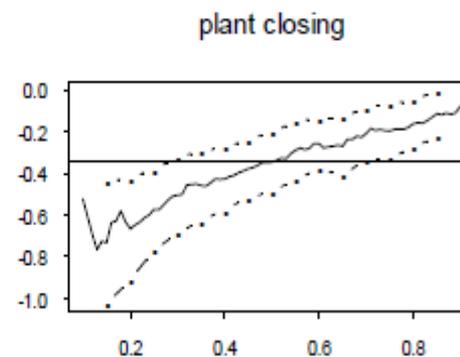
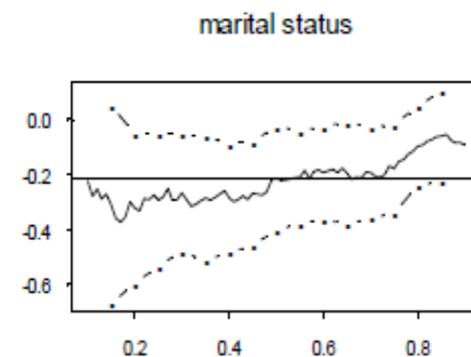
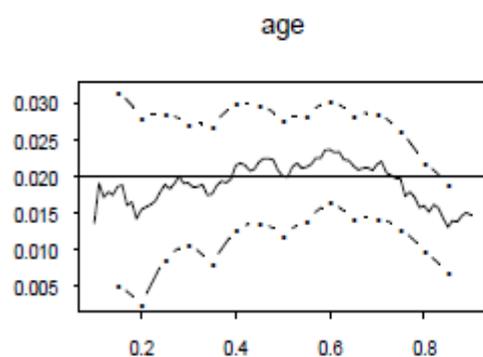
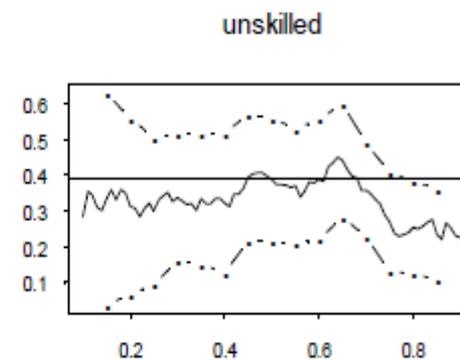
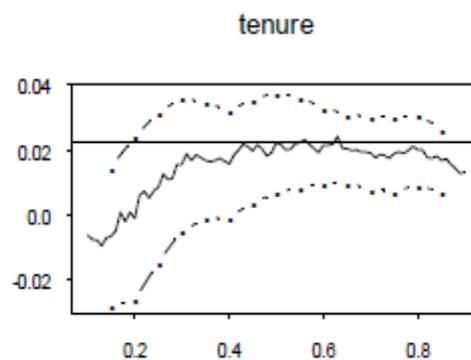
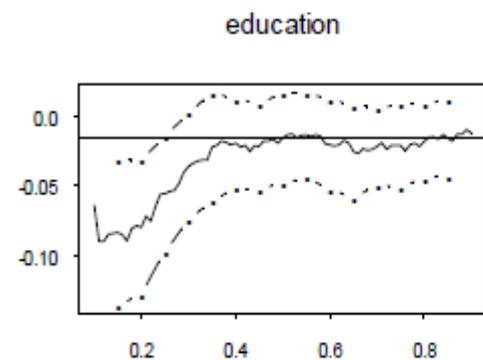
## Revision de literatura

- En cuanto al uso de estimaciones por cuantiles, Koenker y Billias (2001) vuelven a analizar los datos de Pennsylvania Reemployment Bonus Experiments usando esta metodología.
- Los autores encuentran que el bono reduce la duración media del desempleo en alrededor de 10 a 15%, *pero que este efecto esta considerablemente atenuado lejos de la mediana, y esencialmente es insignificante en las colas inferiores y superiores de la distribución*

## Revision de literatura

- Machado y Portugal (2002) usan también esta metodología para analizar la duración del desempleo en Estados Unidos. Ellos encuentran que algunas variables impactan solamente el desempleo de corta duración, mientras que el impacto de otras variables disminuye significativamente en el curso de la duración del desempleo, mientras que el efecto de otras permanece constante.

## Revisión de literatura



## Revisión de literatura

- Fitzenberger y Wilke (2007) calculan un modelo censurado de Box-Cox cuantiles para estudiar la duración del desempleo en Alemania. Los autores resaltan que muchos trabajos empíricos basados en la forma reducida de la duración del desempleo, son consecuencia de un modelo estacionario de búsqueda.
- Modelar hazard rates como funciones de covariadas contemporáneas con coeficientes constantes es insuficiente en ambientes no estacionarios, con agentes que anticipan cambios. De este modo los autores señalan que estimar regresiones por cuantiles es una manera de tomar en cuenta esta no estacionaridad.

## El modelo

$$h(t|x) = \frac{f_i(t|x)}{1 - F_i(t|x)} = \exp(x_i' \hat{\beta}) \lambda_0(t)$$

- En el caso de Weibull,  $\lambda_0(t) = pt^{p-1}$ , el modelo proporcional de riesgo puede ser formulado como un modelo de fallo acelerado:

$$\log(T_i) = x_i' \beta + \varepsilon_i$$

- Donde  $\beta = -p^{-1} \hat{\beta}$  y el termino de error  $\varepsilon_i$  sigue una distribución de valor extremo.

## El modelo

- Este resultado es generalizado de tal manera que cualquier modelo proporcional de riesgo, implica una regresión lineal para una transformación desconocida, tal que:

$$\Lambda_0(t) = \int_0^t \lambda_0(\tilde{t}) d\tilde{t}$$

$$\log(\Lambda_0(T_i)) = x_i' \beta + \varepsilon_i$$

- Dado que  $t = Q_T(\tau|x)$

$$h(\tau|x) \equiv h(Q_T(\tau|x)|x) = \frac{f_{T/x}(Q_T(\tau|x)|x)}{1 - \tau}$$

## El modelo

- $X = [ \text{Jefe del hogar, ingreso per cápita del hogar, la tasa de dependencia, sexo y educación} ]$

$$Q_{\log(T_i)/x_i}(\tau/x_i) = x_i' \beta(\tau)$$

- Tomando en cuenta que la variable de duración es censurada a la derecha se sigue el procedimiento planteado en Pornoy (2003)

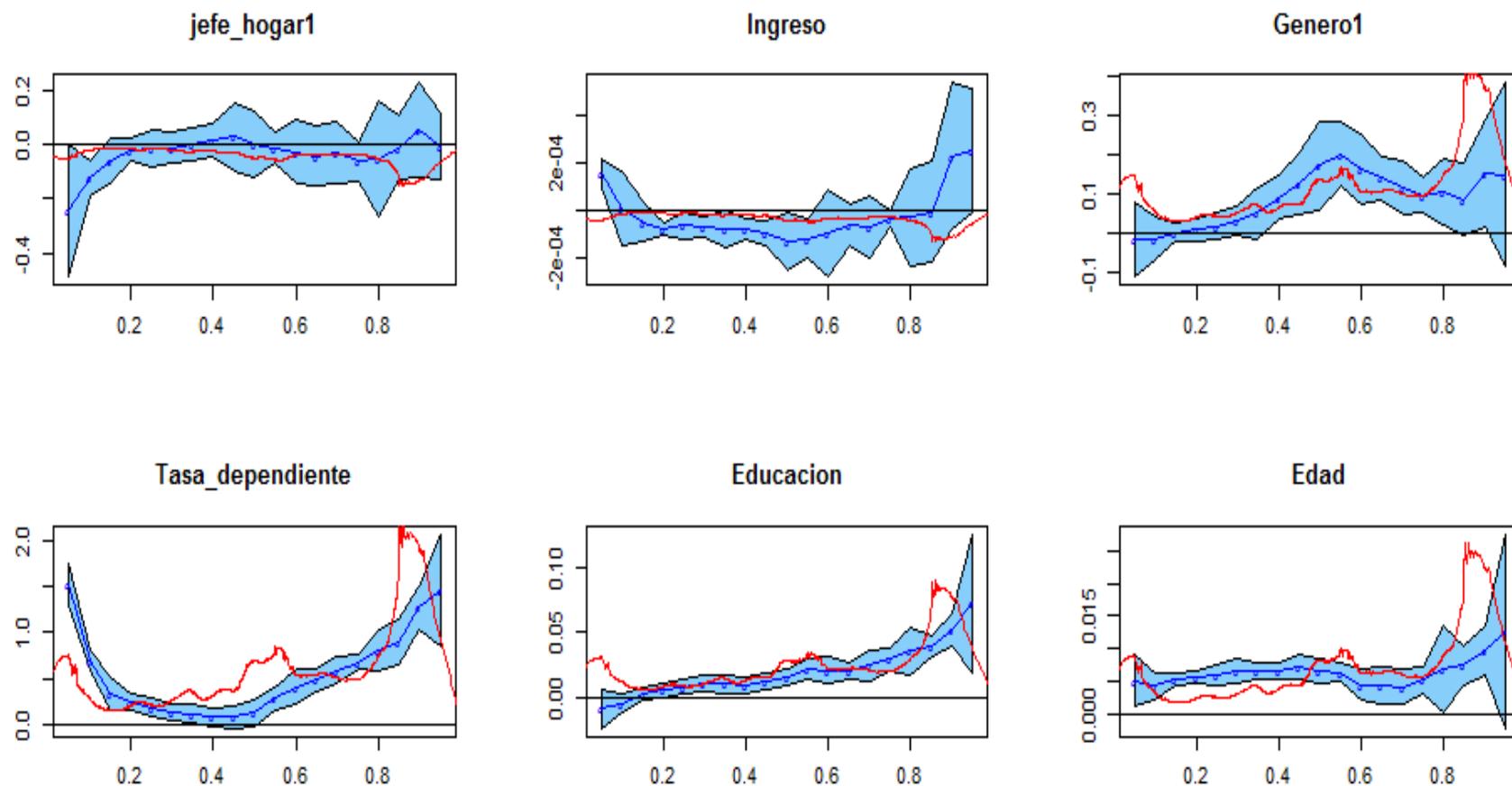
## Los datos

- Los datos se obtienen de la Encuesta Permanente de Empleo que permite obtener información sobre las características de los miembros del hogar, empleo, desempleo e ingresos. Se visitan mensualmente 1600 viviendas seleccionadas por un método aleatorio.
- De esta muestra se ha tomado a los individuos que están desempleados, en los dos periodos, y a los que pasaron de desempleados en el primer periodo a empleados en el segundo. También se han eliminado a los individuos para los que no se ha podido completar la duración de desempleo por falta de información. Con esto, la muestra se redujo a 5003 observaciones.

## Resultados

	Quantile Regression			
	20th	40th	60th	80th
<b>Intercepto</b>	1.64908 (0.0494)	1.87786 (0.07205)	1.98308 (0.12757)	1.91927 (0.1778)
<b>Jefe de Hogar (Jefe=1)</b>	-0.0168 (0.02152)	0.01945 (0.03323)	-0.0255 (0.04374)	-0.0508 (0.08016)
<b>Ingreso</b>	-0.0008 (0.00002)	-0.0008 (0.00002)	-0.0001 (0.00007)	-0.0003 (0.00009)
<b>Genero (Hombre=1)</b>	0.00873 (0.01386)	0.08989 (0.02028)	0.16212 (0.04819)	0.10445 (0.05063)
<b>Tasa de dependencia</b>	0.23397 (0.04428)	0.08278 (0.04764)	0.41113 (0.09737)	0.80624 (0.09945)
<b>Educacion</b>	0.00578 (0.0031)	0.00909 (0.00448)	0.02098 (0.00558)	0.03583 (0.00739)
<b>Edad</b>	0.00571 (0.00038)	0.00668 (0.00090)	0.00452 (0.00131)	0.0069 (0.00307)

## Resultados



## Teoría sobre los resultados

- **Teoría de búsqueda de empleo.** El modelo se centra en la decisión de un desempleado entre aceptar una oferta de trabajo o continuar buscando una mejor.

$$V^{wait}(w) = w^u + \beta EV^{offer}$$

$$V^{accept}(w) = w + \beta \alpha EV^{offer} + \beta (1 - \alpha) V^{accept}(w)$$

$$w^r = w^u + \varphi(w^r)$$

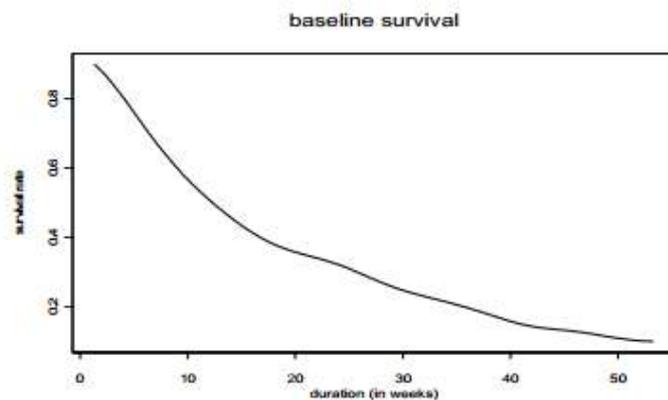
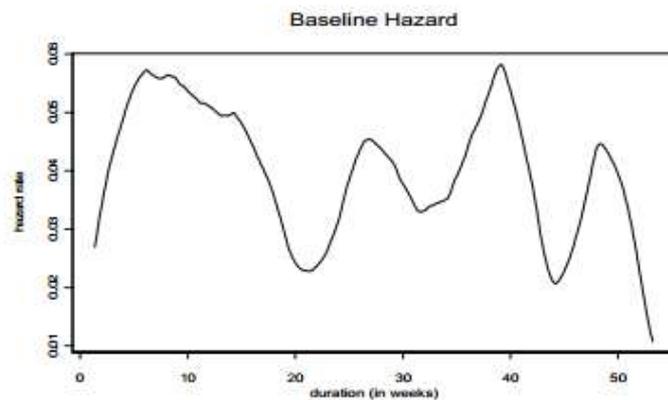
- El lado izquierdo muestra el valor de aceptar un trabajo al salario  $w^r$ , mientras que el lado derecho, nos da el valor de esperar.

## Teoría sobre los resultados

- Fijando un  $w^r$  relativamente alto es menos probable que un trabajador reciba una propuesta aceptable y, en promedio, pasará más tiempo esperando por una oferta.
- La probabilidad de aceptar una oferta, la tasa de aceptación o hazard rate, es igual a la fracción de ofertas mayores o iguales a  $w^r$ .
- Variables como educación, ser hombre parecen aumentar el salario de reserva, y por lo tanto la duración del desempleo.
- El ingreso per cápita del hogar, que podría haber funcionado como un seguro de desempleo,  $w^u$ , no tiene el efecto esperado.

## Pendientes

### Baseline Hazard y survival trabajo de Machado y Portugal (2002)



## Pendientes.

- Test desarrollados Koencker y Machado (1999), y Koencker y Xiao (2002)
- Revisar los datos para tener en cuenta posibles problemas en el proceso de eliminar observaciones.