

# Nonparametric Estimation of Expected Shortfall

ON X C N

*Io St te niversity*

**ABSTRACT** The expected shortfall of a random variable is a popular risk measure. In this paper, we consider the nonparametric estimation of the expected shortfall of a random variable. One of the main results of this paper is that the expected shortfall of a random variable can be estimated by the empirical expected shortfall. The empirical expected shortfall is a natural estimator of the expected shortfall. In this paper, we show that the empirical expected shortfall is a consistent estimator of the expected shortfall. Moreover, we show that the empirical expected shortfall is a uniformly consistent estimator of the expected shortfall. Finally, we show that the empirical expected shortfall is a uniformly consistent estimator of the expected shortfall.

# 1. INTRODUCTION

Let  $\{X_t\}_{t=1}^n$  be a sequence of independent random variables with common distribution  $F$ . Let  $\{Y_t\}_{t=1}^n$  be a sequence of random variables such that  $Y_t = -\log X_t/X_{t-1}$  and  $\{Y_t\}_{t=1}^n$  is independent of  $\{X_t\}_{t=1}^n$ . Let  $\nu_p = \inf\{\mu \in F : \mu \geq -p\}$ .

Let  $\{X_t\}_{t=1}^n$  be a sequence of independent random variables with common distribution  $F$ . Let  $\{Y_t\}_{t=1}^n$  be a sequence of random variables such that  $Y_t = -\log X_t/X_{t-1}$  and  $\{Y_t\}_{t=1}^n$  is independent of  $\{X_t\}_{t=1}^n$ . Let  $\nu_p = \inf\{\mu \in F : \mu \geq -p\}$ .

$$\nu_p = \inf\{\mu \in F : \mu \geq -p\} \tag{1}$$

Let  $\nu_p = \inf\{\mu \in F : \mu \geq -p\}$ . Let  $\nu_p = \inf\{\mu \in F : \mu \geq -p\}$ .



nonparametric regression model. The endogenous variable  $Y_t$  is assumed to be a function of the unobserved variables  $X_t$  and the error term  $\epsilon_t$ . The error term  $\epsilon_t$  is assumed to be independent of  $X_t$  and  $Y_t$ . The nonparametric regression model is defined as follows:

$$Y_t = f(X_t) + \epsilon_t$$

where  $f(\cdot)$  is an unknown function. The nonparametric regression estimator is defined as follows:

$$\hat{f}_n(x) = \frac{1}{n} \sum_{t=1}^n Y_t I(x - X_t) / \frac{1}{n} \sum_{t=1}^n I(x - X_t)$$

where  $I(\cdot)$  is the indicator function. The nonparametric regression estimator is consistent under certain conditions. A detailed proof is given in the appendix.

## 2. NONPARAMETRIC ESTIMATORS

A nonparametric regression estimator of the regression function  $f(\cdot)$  is defined as follows:

$$\hat{f}_n(x) = \frac{Y_{(n(1-p)+1)}}{n} \quad \text{if } x \leq \nu_p \text{ and } Y_{(r)} \quad \text{if } x > \nu_p$$

where  $\nu_p$  is the  $p$ -th quantile of  $\{Y_t\}_{t=1}^n$  and  $Y_{(r)}$  is the  $r$ -th order statistic of  $\{Y_t\}_{t=1}^n$ . Hence,

$$E_p \left[ \frac{\sum_{t=1}^n Y_t I(X_t \geq \nu_p)}{\sum_{t=1}^n I(X_t \geq \nu_p)} \right] = np \quad \text{and} \quad \frac{1}{n} \sum_{t=1}^n Y_t I(X_t \geq \nu_p) \quad (4)$$

where  $I(\cdot)$  is the indicator function. Let  $K$  be a kernel function and  $G_h(\cdot)$  be a kernel density function. Let  $G_h(x) = \int_{-\infty}^{\infty} K(y) dy$  and  $G_h(x) = G(x)/h$  where  $G(\cdot)$  is a positive function and  $h$  is a bandwidth. The nonparametric regression estimator is defined as follows:

$$S_h(x) = \frac{1}{n} \sum_{t=1}^n G_h(x - Y_t)$$

$$S_h(x) = \frac{1}{n} \sum_{t=1}^n G_h(x - Y_t) \quad (4)$$

A kernel density estimator of  $\nu_p$  is denoted by  $\hat{\nu}_{p,h}$  where  $\hat{\nu}_{p,h}$  is a function of  $S_h(x)$ . By the central limit theorem and  $\nu_p$  is a continuous function of  $G_h$  and  $\hat{\nu}_{p,h}$  is a consistent estimator of  $\nu_p$ . The nonparametric regression estimator is defined as follows:

$$\hat{f}_{p,h}(x) = np \frac{1}{n} \sum_{t=1}^n Y_t G_h(x - \hat{\nu}_{p,h}) - Y_t \quad (4)$$

---

<sup>2</sup>Its statistical properties and how to obtain the standard errors are considered in Chen and Tang (2005). See also Cai (2002) and Fan and Gu (2003) for kernel estimation.







The ode for  $Y_t$  is

n

the energy density of the electric field

$p \leq n/q$  and  $p^2 q \leq C p$ . Applying the Cauchy-Schwarz inequality and the fact that  $\sum_{l=[jp]+1}^{(j+1)p} X_l$  is a sum of independent random variables, we have

$$\begin{aligned}
 & P\{|F_n - \nu_p - \epsilon_n| > C_1 \epsilon_n\} \\
 & \leq P\left\{-\frac{C_1^2 \epsilon_n^2 q}{\sigma^2} \leq \left(\frac{1}{C_1 \epsilon_n}\right)^2 q \alpha\left\{\frac{n}{q}\right\}\right\} \quad \text{A 4}
 \end{aligned}$$

$$\begin{aligned}
 & \leq P\left\{-\frac{C_1^2 \epsilon_n^2 q}{\sigma^2} \leq -C_2 \epsilon_n q\right\} \quad \text{A 4} \\
 & \leq P\{-C_2 \epsilon_n q \leq -C_2 \epsilon_n q\} = 0.
 \end{aligned}$$

Since  $C_2 > 0$  and  $n \epsilon_n^2 \rightarrow \infty$ , we have  $n q \epsilon_n \rightarrow \infty$ . This implies that the probability of the event  $\{|F_n - \nu_p - \epsilon_n| > C_1 \epsilon_n\}$  goes to zero as  $n \rightarrow \infty$ .

$$\left\{\left(\frac{1}{C_1 \epsilon_n}\right)^2 q \alpha\left\{\frac{n}{q}\right\}\right\} \leq C \epsilon_n^{-1/2} q \rho^{[n^{1/2} \log^{-1}(n)/2]} \quad \text{A 4}$$

This completes the proof of Lemma 4.  $\square$

Lemma 4. Under the conditions of Lemma 4, for any  $\kappa > 0$ ,

$$n^{-1} \sum_{t=1}^n |Y_t - \nu_p| \mathbb{I}\{Y_t \geq \nu_p\} - \mathbb{I}\{Y_t \geq \nu_p\} = o_p(n^{-3/4+\kappa}).$$

Proof. Let  $W_t = Y_t - \nu_p \mathbb{I}\{Y_t \geq \nu_p\} - \mathbb{I}\{Y_t \geq \nu_p\}$ . Then  $E W_t = 0$  and  $E W_t^2 = \sigma^2 \mathbb{I}\{Y_t \geq \nu_p\} + \mathbb{I}\{Y_t < \nu_p\}$ .

$$\begin{aligned}
 I_{t1} &= E\{Y_t - \nu_p \mathbb{I}\{\nu_p \leq Y_t < \nu_p\} \mathbb{I}\{\nu_p > \nu_p\}\} \quad \text{nd} \\
 I_{t2} &= E\{Y_t - \nu_p \mathbb{I}\{\nu_p \leq Y_t < \nu_p\} \mathbb{I}\{\nu_p < \nu_p\}\}.
 \end{aligned}$$

We have  $E I_{t1} = I_{t11} - I_{t12}$  and  $E I_{t2} = I_{t21} - I_{t22}$ . Let  $a \in \mathbb{R}$ ,  $a > 0$  and  $\eta > 0$ .

$$\begin{aligned}
 I_{t11} &= E\{Y_t - \nu_p \mathbb{I}\{\nu_p \leq Y_t < \nu_p\} \mathbb{I}\{\nu_p \geq \nu_p\}\} = n^{-a} \eta, \\
 I_{t12} &= E\{Y_t - \nu_p \mathbb{I}\{\nu_p \leq Y_t < \nu_p\} \mathbb{I}\{\nu_p < \nu_p\}\} = n^{-a} \eta, \\
 I_{t21} &= E\{Y_t - \nu_p \mathbb{I}\{\nu_p > Y_t \geq \nu_p\} \mathbb{I}\{\nu_p \leq \nu_p\}\} \quad \text{nd} \\
 I_{t22} &= E\{Y_t - \nu_p \mathbb{I}\{\nu_p > Y_t \geq \nu_p\} \mathbb{I}\{\nu_p > \nu_p\}\} = n^{-a} \eta.
 \end{aligned}$$

Applying the Cauchy-Schwarz inequality, we have

$$|I_{tk1}| \leq \sqrt{E \nu_p - \nu_p^2 P\{|\nu_p - \nu_p| \geq n^{-a} \eta\}}$$

en Lee e nd e f c  $E \nu_p - \nu_p^2 = O(n^{-1})$  p y

$$I_{tk1} \rightarrow \text{e ponen } \gamma \text{ y f } \gamma \quad A 4$$

o e  $I_{t12}$  e no  $|I_{t12}| \leq E\{Y_t - \nu_p I_{\nu_p \leq Y_t < \nu_p} n^{-a} \eta\}$ . e n

$$I_{t12} \leq \int_{\nu_p}^{\nu_p + n^{-a}\eta} dz f(z) dz = O(n^{-2a})$$

en Lee e c y e pp o e c n o  $I_{t22} = O(n^{-2a})$  e e nd A 4

e n y o o a  $- / \gamma$  e e  $\gamma >$  y

$$E W_t = O(n^{-1+\kappa}) \quad A 4$$

fo n y p o e  $\kappa$  n n e p e

$$E\left[n^{-1} \sum Y_t - \nu_p \{I_{Y_t \geq \nu_p} - I_{Y_t \geq \nu_p}\}\right] = O(n^{-1+\kappa}) \quad A 4$$

e no con de  $Var W_t = O(a \in , / 4$

$$\begin{aligned} E W_t^2 &= E\left[Y_t - \nu_p \{I_{Y_t \geq \nu_p} - I_{Y_t \geq \nu_p} I_{Y_t \geq \nu_p} - I_{Y_t \geq \nu_p}\}\right] \\ &= E\left[Y_t - \nu_p \{I_{\nu_p > Y_t \geq \nu_p} - I_{\nu_p > Y_t \geq \nu_p}\}\right] \\ &= E\left[Y_t - \nu_p I_{\nu_p \leq Y_t < \nu_p} \{I_{\nu_p \geq \nu_p - n^{-a}\eta} - I_{\nu_p < \nu_p - n^{-a}\eta} a \eta\}\right] \\ &= E\left[Y_t - \nu_p I_{\nu_p > Y_t \geq \nu_p} \{I_{\nu_p \geq \nu_p - n^{-a}\eta} - I_{\nu_p < \nu_p - n^{-a}\eta}\}\right]. \end{aligned}$$

No e

$$E\{I_{\nu_p \leq Y_t < \nu_p} I_{\nu_p \leq \nu_p - n^{-a}\eta}\} \leq P\{|\nu_p - \nu_p| \geq n^{-a}\eta\} \text{ nd}$$

$$E\{I_{\nu_p > Y_t \geq \nu_p} I_{\nu_p > \nu_p - n^{-a}\eta}\} \leq P\{|\nu_p - \nu_p| \geq n^{-a}\eta\}$$

con e e o ze o e ponen  $\gamma$  y f  $\gamma$  p ed y Lee App y n C y.

z neq x e

$$E\{Y_t - \nu_p I_{\nu_p \leq Y_t < \nu_p} I_{\nu_p \leq \nu_p - n^{-a}\eta}\} \text{ nd}$$

$$E\{Y_t - \nu_p I_{\nu_p > Y_t \geq \nu_p} I_{\nu_p \geq \nu_p - n^{-a}\eta}\}$$





e e . q e e p o y n e o

By Hölder's inequality,  $\|n^{-a} f\|_p \leq \|f\|_q$  for  $a \in [0, 1]$  and  $k' \leq n^c$ ,  $c \in [0, 1]$ ,  $-a \leq c$ . Hence, by the Hölder inequality and the definition of  $A$ , we conclude that for  $n \rightarrow \infty$  we have

$$\|S_{n,1}\|_p \xrightarrow{p} \|S_{n,1}\|_q \quad \text{as } n \rightarrow \infty. \quad \square$$

Let us define  $S_{n,1} = n^{-1/2} \sum_{j=1}^r W_{j,n}$ .  $\square$

By the central limit theorem, we have  $S_{n,1} \xrightarrow{d} N(0, \Sigma)$  and the convergence of  $W_{j,n}$

$\square$

Le  $\eta = E\{p^{-1}Y_t K_h(Y_t - \nu_p)\} = p^{-1} \int \nu_p - \nu_p K(y) f(\nu_p - y) dy = p^{-1} \nu_p f(\nu_p) = O(p^2)$   
 y en consecuencia  $\alpha_n = O(p^2)$  en  $n \rightarrow \infty$   
 y  $Cov\{np^{-1} \sum Y_t K_h(\nu_p - Y_t), \nu_{p,h} - \nu_p\} = O(n^{-1})$  en consecuencia  $A = O(n^{-1})$

$$E\{np^{-1} \sum Y_t K_h(\nu_p - Y_t)\} = \eta E(\nu_{p,h} - \nu_p) = O(n^{-1})$$

$$= -\frac{1}{2} p^{-1} \nu_p f'(\nu_p) \sigma_K^2 = o(p^2) = O(n^{-1}) \quad A = O(n^{-1})$$

Consecuentemente  $A = O(n^{-1})$  y  $A = O(n^{-1})$

$$E(\nu_{p,h}) = \nu_p + \frac{1}{2} p^{-1} \sigma_K^2 f''(\nu_p) = o(p^2) = O(n^{-1})$$

Por lo tanto  $\nu_{p,h} = \nu_p + o(p^2)$  en  $n \rightarrow \infty$

en consecuencia  $\nu_{p,h} = \nu_p + o(p^2)$  en  $n \rightarrow \infty$ .  
 Le  $A_1 = np^{-1} \sum_{t=1}^n \{Y_t G_h(\nu_p - Y_t) - Y_t K_h(\nu_p - Y_t) - \nu_{p,h} - \nu_p\}$  es una variable aleatoria con  $E(A_1) = 0$

en

$$Var(A_1) = Var\{np^{-1} \sum Y_t G_h(\nu_p - Y_t)\} + Var\{\eta(\nu_{p,h} - \nu_p)\} - Cov\{np^{-1} \sum Y_t G_h(\nu_p - Y_t), \eta(\nu_{p,h} - \nu_p)\} \quad A = O(n^{-1})$$

Por lo tanto

$$Var\{np^{-1} \sum Y_t G_h(\nu_p - Y_t)\} = n^{-1} p^{-2} \left[ Var\{Y_t G_h(\nu_p - Y_t)\} - \sum_{k=1}^{n-1} \frac{k}{n} Cov\{Y_1 G_h(\nu_p - Y_1), Y_{k+1} G_h(\nu_p - Y_{k+1})\} \right]$$

Le  $\mathcal{C}_K = \int_{-\infty}^{\infty} y K(y) dy = \int_{-\infty}^{\infty} K(y) dy = 0$  y en consecuencia

$$Var\{Y_t G_h(\nu_p - Y_t)\} = \int z^2 G_h^2(\nu_p - z) f(z) dz - p^2 \nu_p^2 = O(p^2)$$

$$= \int_{-\infty}^{\infty} K(y) dy \left[ \int_{-\infty}^u K(y) dy \left\{ \int_{\nu_p}^{\infty} z^2 f(z) dz + \int_{\nu_p - hu}^{\nu_p} z^2 f(z) dz \right\} \right. \\ \left. - \int_u^{\infty} K(y) dy \left\{ \int_{\nu_p}^{\infty} z^2 f(z) dz + \int_{\nu_p - hu}^{\nu_p} z^2 f(z) dz \right\} \right] - p^2 \nu_p^2 = O(p^2)$$

$$Var\{Y_t I(Y_t \geq \nu_p)\} = \nu_p^2 f(\nu_p) \mathcal{C}_K = O(p^2) \quad A = O(n^{-1})$$

Por lo tanto  $A = O(n^{-1})$  y  $A = O(n^{-1})$

$$Var\{np^{-1} \sum Y_t G_h(\nu_p - Y_t)\} = p^{-2} Var\{\phi_1(\nu_p)\} = n^{-1} \nu_p^2 f(\nu_p) \mathcal{C}_K = o(n^{-1}) \quad A = O(n^{-1})$$

The second term on the right side of A 4 is

$$\begin{aligned} & \text{Var}\{\eta \nu_{p,h} - \nu_p\} = \eta - \eta^2 \nu_{p,h} - \nu_p \\ & \eta^2 \text{Var}\{\nu_{p,h}\} = \eta \text{Cov}\{\nu_{p,h}, \eta - \eta^2 \nu_{p,h} - \nu_p\} + \text{Var}\{\eta - \eta^2 \nu_{p,h} - \nu_p\}. \end{aligned}$$

By using the following lemma,  $\eta = p^{-1} \nu_p f(\nu_p) + O_p(n^{-2})$

$$\eta^2 \text{Var}\{\nu_{p,h}\} = p^{-2} \nu_p^2 \text{Var}\{n^{-1} \sum_{t=1}^n I(Y_t > \nu_p)\} - p^{-2} n^{-1} b \nu_p^2 f(\nu_p) c_K + o(n^{-1} b) \quad \text{A 4}$$

By the inequality  $E|Y_t - \nu_p| \leq C n^{-2}$  and  $E|\eta - \eta^2 \nu_{p,h} - \nu_p| = O(n^{-2-3})$

$$E|\nu_{p,h} - \nu_p| \leq C n^{-2} \quad \text{and} \quad E|\eta - \eta^2 \nu_{p,h} - \nu_p| = O(n^{-2-3})$$

Applying the Cauchy-Schwarz inequality and Lemma

$$\text{Var}\{\eta - \eta^2 \nu_{p,h} - \nu_p\} = O(n^{-2-3/2}) = o(n^{-1}) \quad \text{and} \quad \text{A 4}$$

$$\text{Cov}\{\eta \nu_{p,h} - \nu_p, p^{-1} \eta - \eta^2 \nu_{p,h} - \nu_p\} = o(n^{-1}) \quad \text{A 4}$$

Consequently, A 4, A 4 and A 4

$$\text{Var}\{\eta \nu_{p,h} - \nu_p\} = p^{-2} \nu_p^2 \text{Var}\{\phi_2(\nu_p)\} - p^{-2} n^{-1} \nu_p^2 f(\nu_p) c_K + o(n^{-1}) \quad \text{A 4}$$

By Lemma 2.1, the covariance term on the right side of A 4 is

$$\begin{aligned} & \text{Cov}\left\{np^{-1} \sum_{t=1}^n Y_t G_h(\nu_p) - Y_t, \eta \nu_{p,h} - \nu_p\right\} \\ & \text{Cov}\left\{np^{-1} \sum_{t=1}^n Y_t G_h(\nu_p) - Y_t, \eta f^{-1}(\nu_p) n^{-1} \sum_{t=1}^n G_h(\nu_p) - Y_t\right\} = o(n^{-1}) \\ & np^2 p^{-1} \nu_p \left[ \text{Cov}\{Y_t G_h(\nu_p) - Y_t, G_h(\nu_p) - Y_t\} \right. \\ & \left. - \sum_{k=1}^{n-1} \frac{k}{n} \text{Cov}\{Y_1 G_h(\nu_p) - Y_1, G_h(\nu_p) - Y_{k+1}\} \right] = o(n^{-1}) \end{aligned}$$

Since  $\text{Cov}\{Y_t G_h(\nu_p) - Y_t, G_h(\nu_p) - Y_t\} = p^{-1} \nu_p - p \nu_p - \nu_p f(\nu_p) c_K + o(1)$

$$\text{Cov}\left\{np^{-1} \sum_{t=1}^n Y_t G_h(\nu_p) - Y_t, np^{-1} \sum_{i=1}^n Y_i K_h(\nu_p) - Y_t \nu_{p,h} - \nu_p\right\} = \text{A 4}$$

$$n^{-1} p^{-2} \nu_p \text{Cov}\{\phi_1(\nu_p), \phi_2(\nu_p)\} - n^{-1} p^{-2} \nu_p^2 f(\nu_p) c_K + o(1)$$

$O(n^{-1})$  variance of  $\hat{\mu}_p$  and  $O(n^{-1})$  variance of  $\hat{\sigma}_0^2$

$$\text{Var}(\hat{\mu}_p) = p^{-1}n^{-1}\sigma_0^2 \quad \text{and} \quad \text{Var}(\hat{\sigma}_0^2) = O(n^{-1})$$

Theorem 4

Let  $\nu_{p,h}$  be the density of  $\nu_{p,h}$  and  $\nu_{p,h}$  be the density of  $\nu_{p,h}$ .

### ~~CONFIDENTIAL~~

Theorem 4.1. Let  $\nu_{p,h}$  be the density of  $\nu_{p,h}$ .

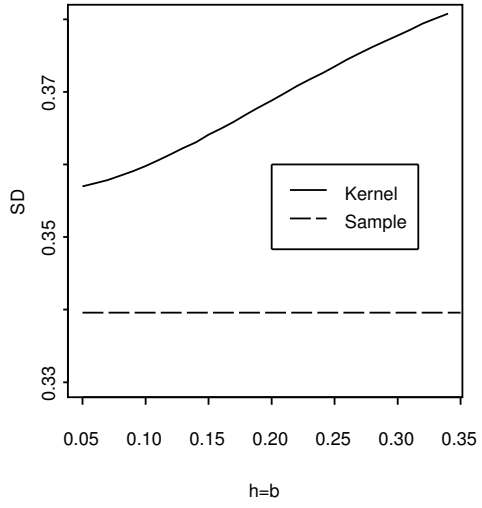
Theorem 4.2. Let  $\nu_{p,h}$  be the density of  $\nu_{p,h}$ .

Theorem 4.3. Let  $\nu_{p,h}$  be the density of  $\nu_{p,h}$ .

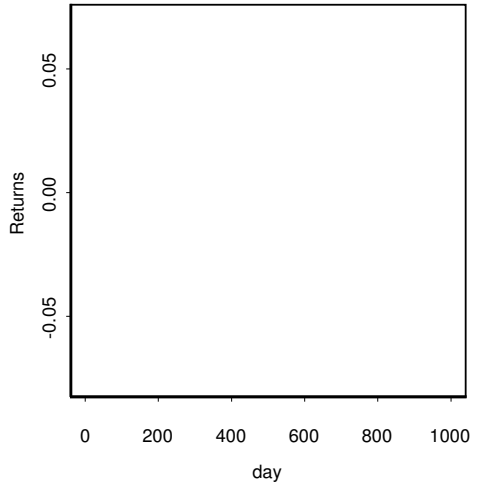
Theorem 4.4. Let  $\nu_{p,h}$  be the density of  $\nu_{p,h}$ .

- ey and McNe A 4 and e pec ed of n po fo of dependen  
 c ed concep nd p c c n oun of B nking nd Fin nce
- enon C o e e nd C L edo 4 c o y Mode  
 dden M o Mode nd c App c on Bernoulli
- o e o C c e O nd L en 4 en x n y of e  
 oun of e piric n nce
- o on 4 t Risk nd on Ne Yo Mc
- Le d e e M L nd en nd oo zen 4 E tre es nd Re ved Properties  
 of R ndo Sequences nd Processes p n e Be n
- M y nd e e D 4 Nonp e e c e on nd den c on of non  
 ne e e e Econo etric Theory 11
- e D nd e M 4 St tistic An ysis of E tre e es ith pp i  
 c tions to insur nce hydrogy nd other eds nd ed Bo on B
- e
- on M 4 A cen e e e nd e a n cond on Proc N t  
 Ac d Sc S A 42
- c e O 4 Nonp e e c e on of cond on e pec ed of C en  
 e e e e o
- c e O 4 Nonp e e c e on nd en x n y of e pec ed of  
 M the tic Fin nce 14
- e e nd M on 4 e ne q n e e oun of A eric n  
 St tistic Associ tion 5
- Yo oy 4 Mo en o nd fo on y e n q ence Pro i ty Theory  
 nd Re ved Fie ds 52
- Ye 4 e B d e p e n on of e p e q n e fo q ence of  
 on y e n nd e e St tistics nd Pro i ty Letters 24

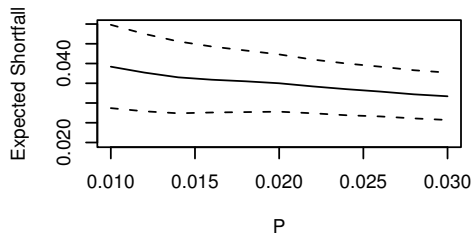
(a) ES Estimates



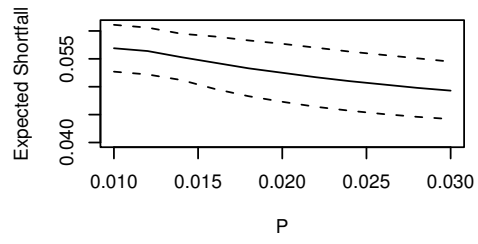




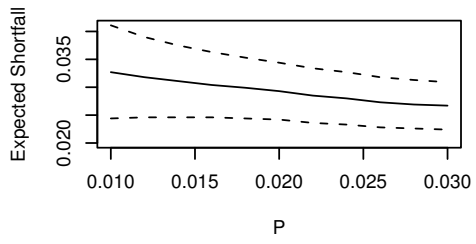
DO ON  $\lambda^2$   $1^2$



CAC  $\lambda^2$   $1^2$



DO ON  $\lambda^2$   $2^2$



*T* Esti tes for  $\nu_{0.01}$   $\mu_{0.01}$  nd St nd rd Errors S.E

| Ye | CAC          |         | Do one       |         |
|----|--------------|---------|--------------|---------|
|    | $\nu_{0.01}$ | $\mu_p$ | $\nu_{0.01}$ | $\mu_p$ |
| •  |              |         |              |         |
| •  |              |         |              |         |
| •  |              |         |              |         |