151163A - Financial Econometrics

I. Primer on Financial Econometrics

CHEUNG Ying Lun

Capital University of Economics and Business

Asset Returns

Shanghai Stock Exchange Composite Index

Source: http://stockhtm.finance.qq.com/hqing/zhishu/000001.htm

One-period simple return

If we hold an asset for one period, from $t-1$ to t, the *simple gross return* is given by:

$$1 + R_t = \frac{P_t}{P_{t-1}} \qquad \text{or} \qquad P_t = P_{t-1}(1 + R_t).$$

The one-period *simple net return* or *simple return* is

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{\Delta P_t}{P_{t-1}}.$$

Shanghai Stock Exchange Composite Index

Source: http://stockhtm.finance.qq.com/hqing/zhishu/000001.htm

Multi-period simple return

If we hold an asset for k periods, from $t-k$ to t, the k-period simple gross return is given by:

$$\begin{aligned} 1+R_t[k] = \frac{P_t}{P_{t-k}} &= \frac{P_t}{P_{t-1}}\frac{P_{t-1}}{P_{t-2}}\cdots\frac{P_{t-k+1}}{P_{t-k}} \\ &= (1+R_t)(1+R_{t-1})\ldots(1+R_{t-k+1}) \\ &= \prod_{j=0}^{k-1}(1+R_{t-j}). \end{aligned}$$

Similarly, the k-period net return is

$$R_t[k] = \frac{P_t - P_{t-k}}{P_{t-k}} = \prod_{j=0}^{k-1}(1+R_{t-j}) - 1.$$

Day	P_{t-1}	P_t	ΔP_t	R_t	$R_t[k]$
1	3031.24	3030.75	-0.49	-0.02%	-0.02%
2	3030.75	2979.40	-51.35	-1.69%	-1.71%
3	2979.40	2985.66	6.26	0.21%	-1.50%
4	2985.66	2999.28	13.62	0.46%	-1.05%
5	2999.28	3006.45	7.17	0.24%	-0.82%

Table: One-period and multi-period simple returns

On average, what is the return per period?

Average return

Let $\overline{R}$ be the average return,

$$\begin{aligned} 1 + R_t[k] &= (1 + R_t)(1 + R_{t-1}) \dots (1 + R_{t-k+1}) \\ &= (1 + \overline{R})(1 + \overline{R}) \dots (1 + \overline{R}) \\ &= (1 + \overline{R})^k \end{aligned}$$

which yields

$$\overline{R} = (1 + R_t[k])^{1/k} - 1 = \left[\prod_{j=0}^{k-1} (1 + R_{t-j}) \right]^{1/k} - 1.$$

If each period spans one year, then $\overline{R}$ is also called the *annualized* return.

Suppose you are going to deposit \$10,000 in a bank, which offers you a 10% per annum interest rate and the following compounding scheme:

1. Compounding every year, where the one-year interest rate is 10%;
2. Compounding every 6 months, where the 6-month interest rate is $10\%/2 = 5\%$.

Which one should you choose?

Effect of compounding

Type	No. of payments	Interest rate per period	Total value
Annual	1	10%	$11000.00
Semiannual	2	5%	$11025.00
Quarterly	4	2.5%	$11038.13
Monthly	12	0.833%	$11047.13
Weekly	52	0.192%	$11050.65
Daily	365	0.027%	$11051.56

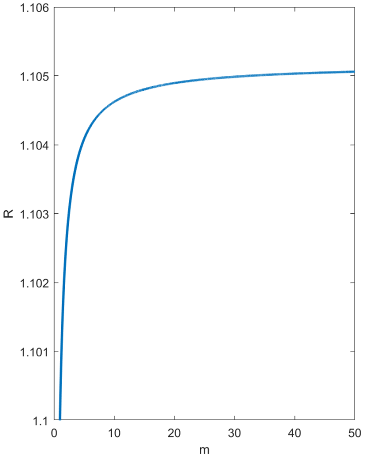
Table: Values of a loan with 10% per annum interest rate

Effect of compounding

In general, if the bank gives interest m times a year, you get

$$\$10,000 \times \left(1 + \frac{10\%}{m}\right)^m.$$

What if $m \to \infty$?



Continuous compounding

Suppose the continuously compounded interest rate is r_t, the simple gross return, or the *effective annual interest rate*, is

$$1 + R_t = \lim_{m\to\infty} \left(1 + \frac{r_t}{m}\right)^m .$$

Taking logarithm, and by L'Hopital's Rule

$$\lim_{m\to\infty} m \ln \left(1 + \frac{r_t}{m}\right) = r_t .$$

Therefore, $1 + R_t = \mathrm{e}^{r_t}$, or $r_t = \ln(1 + R_t)$, where r_t is also called the log return.

Continuously compounded returns

The one-period log return is given by

$$r_t = \ln(1 + R_t) = \ln \frac{P_t}{P_{t-1}} = p_t - p_{t-1}$$

where $p_t = \ln P_t$. The multi-period log return is given by

$$\begin{aligned} r_t[k] &= \ln(1 + R_t[k]) = \ln[(1 + R_t)(1 + R_{t-1}) \dots (1 + R_{t-k+1})] \\ &= \ln(1 + R_t) + \ln(1 + R_{t-1}) + \cdots + \ln(1 + R_{t-k+1}) \\ &= r_t + r_{t-1} + \cdots + r_{t-k+1} \end{aligned}$$

Suppose the log-return is constant $r_t = r$, and the price of an asset at time 0 is P_0, then

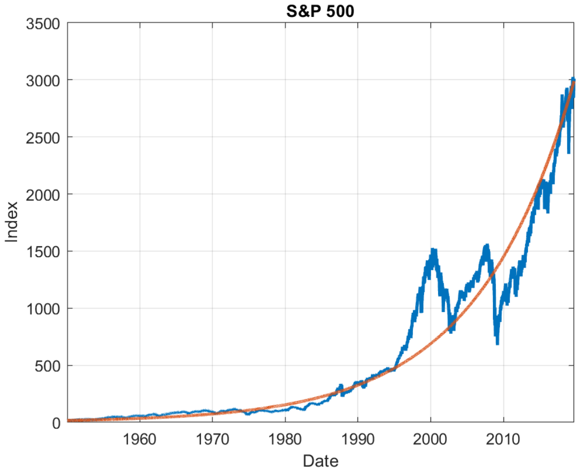
$$r_t[t] = r_t + \cdots + r_1 = t \cdot r.$$

Moreover,

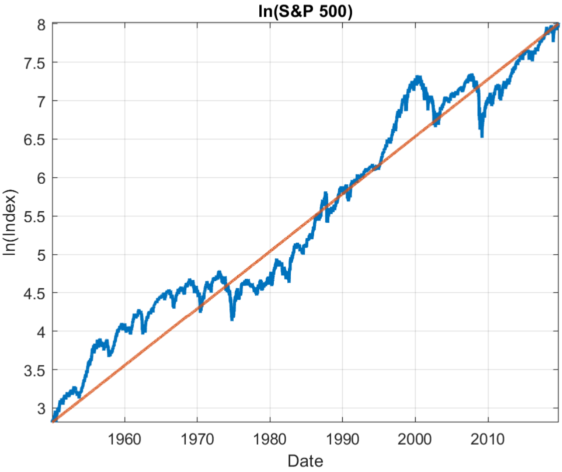
$$\begin{aligned} P_t &= P_0 \cdot (1 + R_1) \ldots (1 + R_t) \\ &= P_0 \cdot \mathrm{e}^r \ldots \mathrm{e}^r \\ &= P_0 \cdot \mathrm{e}^{r \cdot t}. \end{aligned}$$

The asset price will be growing *exponentially*.

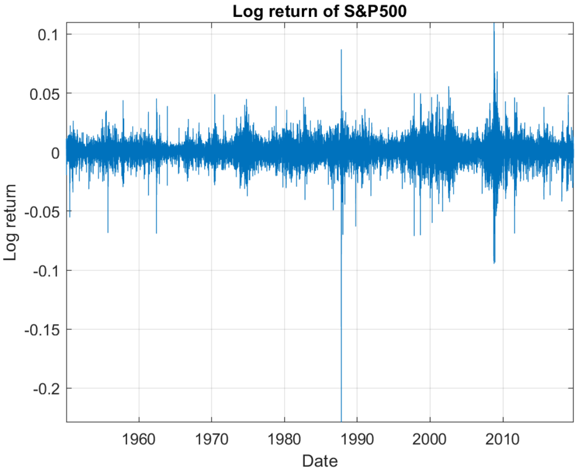
S&P 500 historical values



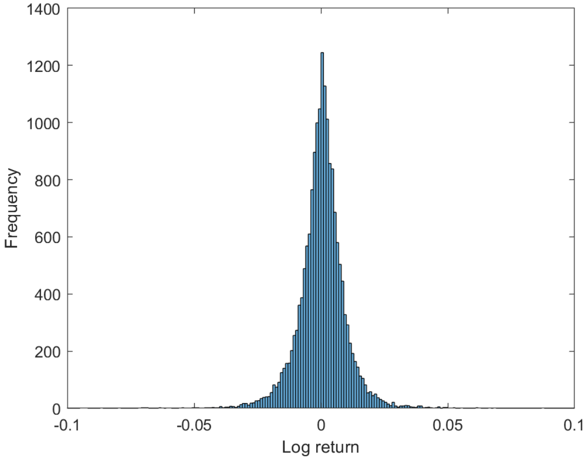
Logarithm of S&P 500



Log returns of S&P 500



Distribution of S&P 500 log returns



Review of Statistical Distributions

Joint density function

Suppose X and Y are two random variables with support $(-\infty, \infty)$, with parameters $\boldsymbol{\theta}$. We define the *joint distribution function* as

$$F_{X,Y}(x, y; \boldsymbol{\theta}) = P(X \leq x, Y \leq y; \boldsymbol{\theta}).$$

If the *joint probability density function* $f_{X,Y}(x, y; \boldsymbol{\theta})$ exists, then

$$F_{X,Y}(x, y; \boldsymbol{\theta}) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(w, z; \boldsymbol{\theta}) dz dw.$$

A simple example: Uniform distribution

Let X, Y be two random variables with a joint density function

$$f_{X,Y}(x, y) = \begin{cases} 1, & \text{if } x \in [0, 1],\ y \in [0, 1] \\ 0, & \text{otherwise.} \end{cases}$$

Find

$$F_{X,Y}(0.5, 0.8) = P(X \leq 0.5, Y \leq 0.8).$$

The *conditional distribution* of X given $Y \leq y$ is given by

$$F_{X|Y\leq y}(x;\boldsymbol{\theta}) = \frac{P(X \leq x, Y \leq y; \boldsymbol{\theta})}{P(Y \leq y; \boldsymbol{\theta})}.$$

The *conditional density* is

$$f_{X|Y}(x;\boldsymbol{\theta}|Y = y) = \frac{f_{X,Y}(x, y; \boldsymbol{\theta})}{f_Y(y; \boldsymbol{\theta})}$$

where the *marginal density function* $f_Y(y;\boldsymbol{\theta})$ is given by

$$f_Y(y;\boldsymbol{\theta}) = \int_{-\infty}^{\infty} f_{X,Y}(x, y; \boldsymbol{\theta})dx.$$

A simple example (continued)

Let X, Y be two random variables with a joint density function

$$f_{X,Y}(x,y) = \begin{cases} 1, & \text{if } x \in [0,1],\ y \in [0,1] \\ 0, & \text{otherwise.} \end{cases}$$

Find

$$F_{X|Y \leq 0.8}(0.5) = P(X \leq 0.5 | Y \leq 0.8).$$

The j-th moment of a random variable X is defined as

$$m_j' = \mathbb{E}\left[X^j\right] = \int_{-\infty}^{\infty} x^j f(x) dx.$$

Let $\mu_X = \mathbb{E}[X] = m_1'$, the j-th centered moment of X is

$$m_j = \mathbb{E}\left[(X - \mu_X)^j\right] = \int_{-\infty}^{\infty} (x - \mu_X)^j f(x) dx.$$

Moments

- The first moment is the *mean*, which measures the (average) location of X.

$$\mu_X = \mathbb{E}[X]$$

The sample mean is

$$\widehat{\mu}_X = \frac{1}{T}\sum_{t=1}^{T} x_t$$

- The second centered moment is the *variance*, which measures the dispersion of X around its mean.

$$\sigma_X^2 = \mathbb{E}\left[(X - \mu_X)^2\right]$$

The sample variance is

$$\widehat{\sigma}_X^2 = \frac{1}{T-1}\sum_{t=1}^{T}(x_t - \widehat{\mu}_X)^2$$

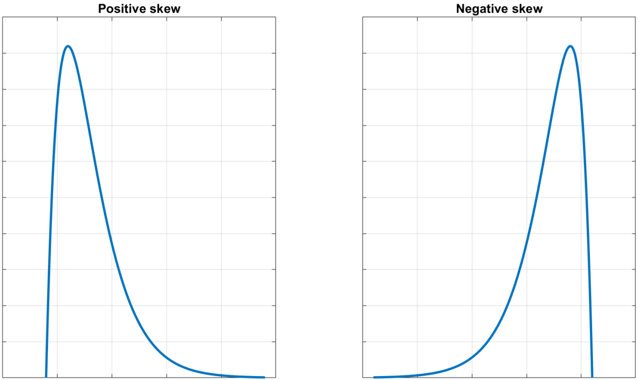
- The third centered moment is *skewness*, which measures the degree of asymmetry in the distribution of X.

$$S(X) = \mathbb{E}\left[\frac{(X - \mu_X)^3}{\sigma_X^3}\right]$$

The sample skewness is

$$\widehat{S}(X) = \frac{1}{(T-1)\widehat{\sigma}_X^3}\sum_{t=1}^{T}(x_t - \widehat{\mu}_X)^3$$

Skewness



Kurtosis

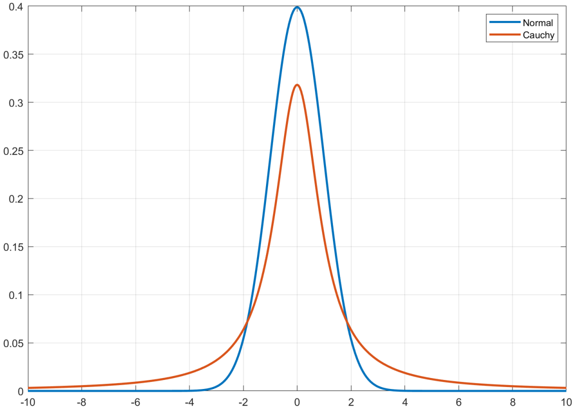
- The fourth centered moment is *kurtosis*, which measures the fatness of the tails of the distribution of X.

$$K(X) = \mathbb{E}\left[\frac{(X-\mu_X)^4}{\sigma_X^4}\right]$$

The sample kurtosis is

$$\widehat{K}(X) = \frac{1}{(T-1)\widehat{\sigma}_X^4}\sum_{t=1}^{T}(x_t - \widehat{\mu}_X)^4$$

Kurtosis



Let X be two random variables with a joint density function

$$f_X(x) = \begin{cases} 1, & \text{if } x \in [0, 1] \\ 0, & \text{otherwise.} \end{cases}$$

Find the mean, variance, skewness and kurtosis of X.

Normal distribution

Suppose $X \sim \mathcal{N}(\mu, \sigma^2)$, then its density function is

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

The population moments are

Mean	Variance	Skewness	Kurtosis
μ	σ^2	0	3

A random variable with normal distribution has skewness 0 and kurtosis 3. Moreover, asymptotically,

$$\widehat{S}(x) \xrightarrow{d} \mathcal{N}\left(0, \frac{6}{T}\right), \qquad \widehat{K}(x) \xrightarrow{d} \mathcal{N}\left(3, \frac{24}{T}\right).$$

The hypothesis of normality can be tested using the t-statistics

$$t_S = \frac{\widehat{S}(x)}{\sqrt{6/T}}, \qquad t_K = \frac{\widehat{K}(x) - 3}{\sqrt{24/T}}.$$

Alternatively, one can also use the Jarque and Bera (JB) test statistic

$$JB(x) = \frac{\widehat{S}(x)^2}{6/T} + \frac{[\widehat{K}(x) - 3]^2}{24/T} \xrightarrow{d} \chi_2^2$$

Summary statistics of S&P 500 returns

	Sample moments	**t-stat**	**95% Critical values**	**Reject H_0**
Mean	0.704	-	-	-
Variance	17.268	-	-	-
Skewness	-0.450	-5.313	(-1.96, 1.96)	✓
Kurtosis	5.157	12.727	(-1.96, 1.96)	✓
JB Statistic	190	-	5.99	✓

Problems of using the normal distribution:

1. The lower bound of the simple return is -1, but the support of the normal distribution has no lower bound.
2. If R_t is normally distributed, then $R_t[k]$ is not normally distributed.
3. Empirical asset returns tend to have positive excess kurtosis.

If we assume $r_t = \ln(1 + R_t) \sim \mathcal{N}(\mu, \sigma^2)$, then we say R_t is log-normally distributed. In this case,

$$\mathbb{E}[R_t] = \mathrm{e}^{\mu + \frac{\sigma^2}{2}} - 1, \qquad \mathrm{var}(R_t) = \mathrm{e}^{2\mu+\sigma^2}[\mathrm{e}^{\sigma^2} - 1]$$

If R_t is log-normally distributed with mean and variance m_1 and m_2, then we can show that

$$\mu = \ln\left(\frac{m_1 + 1}{\sqrt{1 + m_2/(1+m_1)^2}}\right), \qquad \sigma^2 = \ln\left(1 + \frac{m_2}{(1+m_1)^2}\right).$$

Summary statistics of S&P 500 log returns

	Sample moments	**t-stat**	**95% Critical values**	**Reject H_0**
Mean	0.615	-	-	-
Variance	17.484	-	-	-
Skewness	-0.712	-8.401	(-1.96, 1.96)	✓
Kurtosis	5.877	16.982	(-1.96, 1.96)	✓
JB Statistic	359	-	5.99	✓

Advantage of using the log-normal distribution

1. $r_t[k]$ is the sum of normally distributed random variables and is still normally distributed.
2. There is no lower bound for r_t and $R_t = \mathrm{e}^{r_t} - 1 \geq 0$ is still satisfied.

Problem of using the log-normal distribution

1. Empirical asset log returns tend to have positive excess kurtosis.

The log-return r_t follows a scale mixture if $r_t \sim \mathcal{N}(\mu, \sigma^2)$, where σ^2 follows a positive distribution. For example,

$$r_t \sim (1 - X)\mathcal{N}(\mu, \sigma_1^2) + (1 - X)\mathcal{N}(\mu, \sigma_2^2)$$

where X is a Bernoulli random variable such that $P(X = 1) = \alpha$ and $P(X = 0) = 1 - \alpha$, with $0 < \alpha < 1$. Here σ_1^2 is small and σ_2^2 is relatively large.

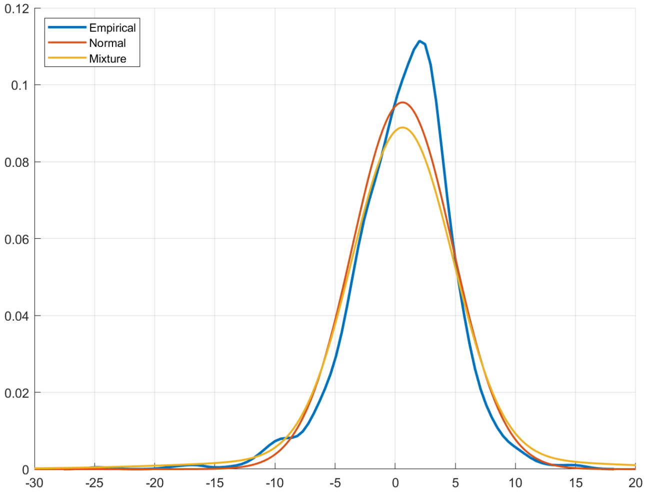
Advantage of using the scale mixture of normal distribution:

1. It maintains the tractability of normal.
2. Higher order moments are still finite.
3. It can capture the excess kurtosis.

Problem of using the scale mixture of normal distribution:

1. It is hard to estimate the mixture parameter α.

Empirical density of S&P 500



VaR and Expected Shortfall

Value at risk (VaR)

What is the potential for loss of an asset with a certain probability?

The VaR is the potential loss that happens with a specified probability. Let ΔV be the change in values of an asset, then VaR is defined as

$$P(\Delta V \geq VaR) = F(VaR)$$

where $F(\cdot)$ is the cumulative distribution function (CDF) of ΔV.

The α-VaR can be obtained as the α-quantile of ΔV, i.e.,

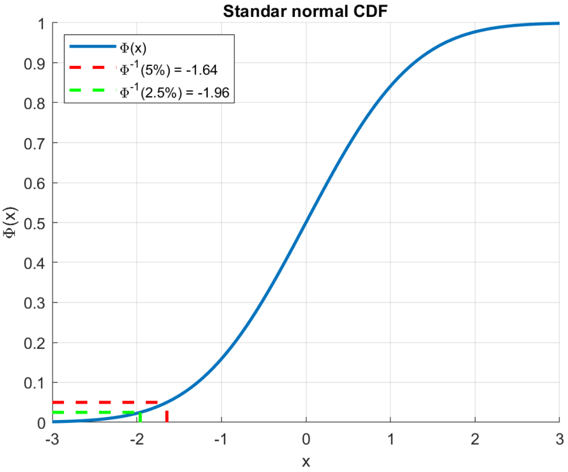
$$VaR_\alpha = \inf\{\Delta V | F(\Delta V) \geq \alpha\}$$

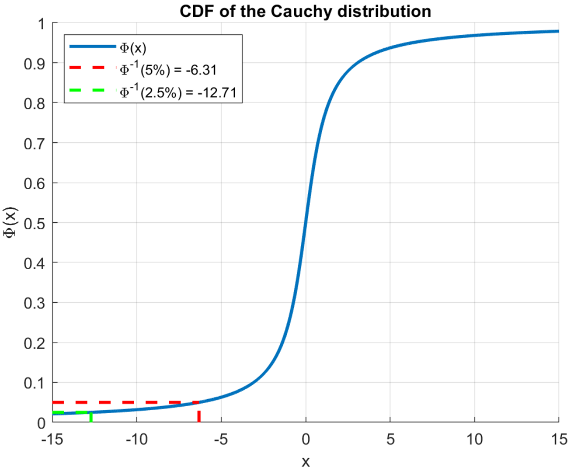
If $\Delta V \sim \mathcal{N}(\mu, \sigma^2)$, then

$$VaR_\alpha = \mu + \sigma \Phi^{-1}(\alpha),$$

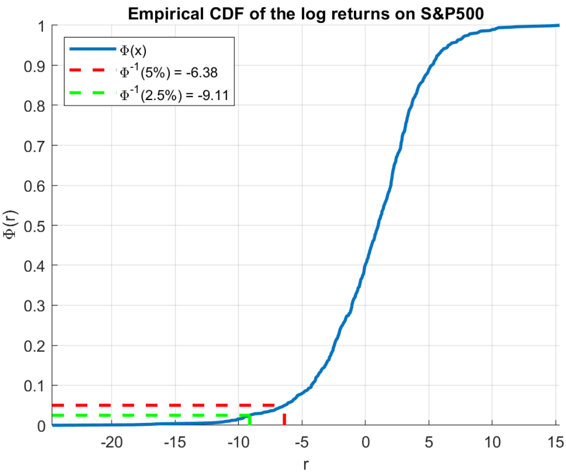
where $\Phi^{-1}(\alpha)$ is the inverse CDF of a standard normal distribution.

Value at risk (VaR)




CDF of the Cauchy distribution
Φ(x)
Φ-1(5%) = -6.31
Φ-1(2.5%) = -12.71
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-15
-10
-5
0
5
10
15
x
Φ(x)

Value at risk (VaR)



Expected shortfall

VaR only tells you that with $\alpha\%$ chance, you can loss *more* than VaR_α, but it does not tell you *how much* you can loss. Consider the following extreme case:

The expected shortfall at $\alpha\%$ level is the expected return on the portfolio/an asset in the worst $\alpha\%$ of cases, i.e.,

$$ES_\alpha(\Delta V) = \mathbb{E}\left[\Delta V | \Delta V \leq VaR_\alpha\right].$$

The expected shortfall is therefore the average loss given that the loss exceeds the VaR. *Importantly, it uses the whole tail of the distribution instead of just a single quantile.*