In [1]:
```
using Gadfly
using Distributions
using Interact
```

---

https://faculty.washington.edu/ezivot/econ424/timeseriesconcepts.pdf (https://faculty.washington.edu/ezivot/econ424/timeseriesconcepts.pdf)

1.1 Stochastic Processes

A stochastic process is a sequence of random variables indexed by time t.

$$\{\ldots, Y_1, Y_2, \ldots, Y_t, Y_{t+1}, \ldots\} = \{Y\}_{t=-\infty}^{T}$$

In modeling time series data, the ordering imposed by the time index is important because we often would like to capture the temporal relationships, if any, between the random variables in the stochastic process. In random sampling from a population, the ordering of the random variables representing the sample does not matter because they are independent.

A realization of a stochastic process with T observations is the sequence of observed data.

$$\{Y_1 = y_1, Y_2 = y_2, \ldots, Y_T = t_T\} = \{Y\}_{t=1}^{T}$$

1.1.1 Stationary Stochastic Processes

Definition *Strict stationarity*

A stochastic process $\{Y_t\}_{t=-\infty}^{\infty}$ is strictly stationary if, for any given finite integer r and for any set of subscripts $t_1, t_2, \ldots, t_r$ the joint distribution of $(Y_{t_1}, Y_{t_2}, \ldots, Y_{t_r})$ depends only on $t_1 - t, t_2 - t, \ldots, t_r - t$ but not on t.

In simple terms, the joint distribution of random variables in a strictly stationary stochastic process is time invariant. For example, the joint distribution of (Y_1, Y_5, Y_7) is the same as the distribution of (Y_{12}, Y_{16}, Y_{18}).

Example *iid sequence*

If $\{Y_t\}_{t=-\infty}^{\infty}$ is an iid sequece, then it is strictly stationary.

Definition *Covariance stationarity*

A stochastic process $\{Y_t\}_{t=1}^{\infty}$ is covariance stationary if:

1. $E[Y_t] = \mu$ does not depend on t.

2. $var(Y_t) = \sigma^2$ does not depend on t.

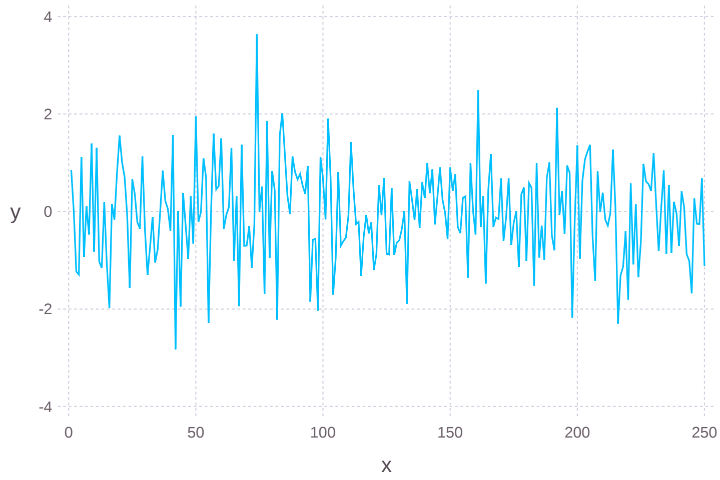
1. $cov(Y_t, Y_{t-j}) = \gamma_j$ exists, is finite, and depends only on j but not on t for $j = 0, 1, 2, \ldots$

If we assume that the stochastic process $\{Y_t\}_{t=-\infty}^{\infty}$ is strictly stationary and that $E[Y_t]$, $var(Y_t)$, and all pairwise covariances exist, then we say that $\{Y_t\}_{t=-\infty}^{\infty}$ is a covariance stationary stochastic process.

Example *Gaussian White Noise*

Let $Y_t \sim \ iid \ N(0, \sigma^2)$. Then $\{Y_t\}_{t=-\infty}^{\infty}$ is called a Gaussian white noise process.

In [2]:
```
plot(x=collect(1:250), y=rand(Normal(0,1), 250), Geom.line)
```
Out[2]:



1.1.2 Non-Stationary Processes

In a covariance stationary stochastic process it is assumed that the means, variances and autocovariances are independent of time. In a non-stationary process, one or more of these assumptions is not true.

Example *Deterministically trending process*

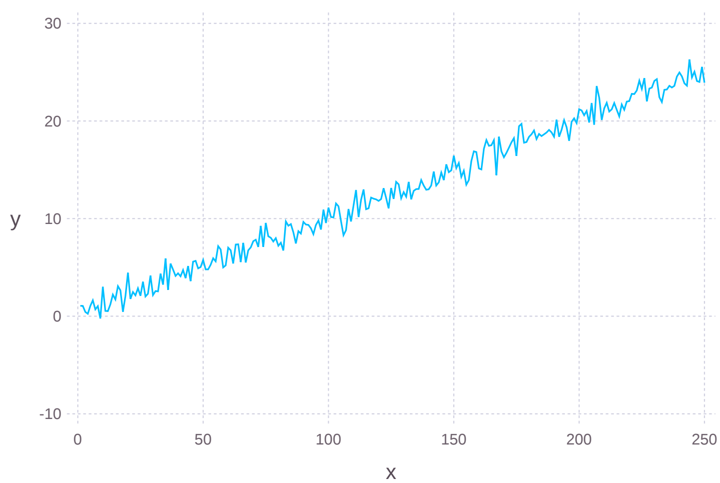
Suppose $\{Y_t\}_{t=0}^{\infty}$ is generated according to the deterministically trending process

$$Y_t = \beta_0 + \beta_1 t + \varepsilon_t \sim WN(0, \sigma_\varepsilon^2) \qquad t = 0, 1, 2, \dots$$

Then $\{Y_t\}_{t=0}^{\infty}$ is nonstationary because the mean of Y_t depends on t:

$$E[Y_t] = \beta_0 + \beta_1 t \qquad \text{depends on } t.$$

In [3]:
```
plot(x=collect(1:250),y=0.1.*collect(1:250)+rand(Normal(0,1), 250), Geom.line)
```
Out[3]:

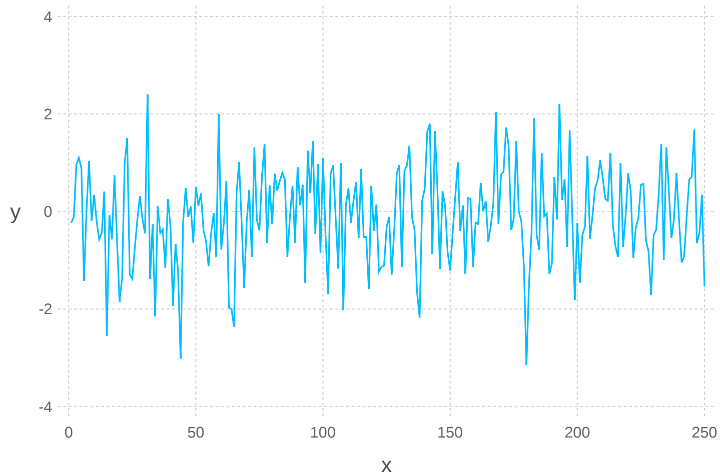


Here the non-stationarity is created by the deterministic trend $Y_t = 0.1t + \varepsilon_t, \varepsilon_t \sim N(0, 1)$ in the data. The non-stationary process $\{Y_t\}_{t=0}^{\infty}$

can be transformed into a stationary process by simply subtracting off the trend:

$$X_t = Y_t - \beta_0 - \beta_1 t = \varepsilon_t \sim \mathrm{WN}(0, \sigma_\varepsilon^2)$$

In [4]:
```
plot(x=collect(1:250),y=(0.1*collect(1:250)+rand(Normal(0,1), 250)) - (0.1*collect(1:250)), Geom.line)
```
Out[4]:



Example *Random walk*

A random walk (RW) process $\{Y_t\}_{t=1}^{\infty}$ is defined as

$$Y_t = Y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathrm{IWN}(0, \sigma_\varepsilon^2), \quad Y_0 \text{ is fixed (non-random).}$$

In [5]:
```
plot(x=collect(1:250),y=cumsum(rand(Normal(0,1), 250)), Geom.line)
```
Out[5]:

1.1.3 Ergodicity

In a strictly stationary or covariance stationary stochastic process no assumption is made about the strength of dependence between random variables in the sequence. However, in many contexts it is reasonable to assume that the strength of dependence between random variables in a stochastic process diminishes the farther apart they become. This diminishing dependence assumption is captured by the concept of ergodicity.

Definition *Ergodicity (intuitive definition)*

Intuitively, a stochastic process $\{Y_t\}_{t=-\infty}^{\infty}$ is ergodic if any two collections of random variables partitioned far apart in the sequence are essentially independent.

The formal definition of ergodicity is highly technical and requires advanced concepts in probability theory. However, the intuitive definition captures the essence of the concept. The stochastic process $\{Y_t\}_{t=-\infty}^{\infty}$ is ergodic if Y_t and Y_{t-j} are essentially independent if j is large enough.

Example *White noise processes*

If $\{Y_t\}_{t=-\infty}^{\infty}$ is GWN then it is both covariance stationary and ergodic.

The different flavors of **white noise processes** are not very interesting because they **do not allow any linear dependence** between the observations in the series. The following sections describe some simple covariance stationary and ergodic time series models that allow for different patterns of time dependence captured by autocorrelations.

1.2 Moving Average Processes

Moving average models are simple covariance stationary and ergodic time series models that can capture a wide variety of autocorrelation patterns.

1.2.1 MA(1) Model

Suppose you want to create a covariance stationary and ergodic stochastic process $\{Y_t\}_{t=-\infty}^{\infty}$ in which Y_t and Y_{t-1} are correlated but Y_t and Y_{t-j} are not correlated for $j > 1$. That is, the time dependence in the process only lasts for one period. Such a process can be created using the first order moving average ($\mathrm{MA}(1)$) model:

$$Y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}, \quad -1 < \theta < 1, \quad \varepsilon_t \sim iid\, N(0, \sigma_\varepsilon^2)$$

.

The moving average parameter θ determines the sign and magnitude of the correlation between Y_t and Y_{t_1}. Clearly, if $\theta = 0$ then $Y_t = \mu + \varepsilon_t$ so that $\{Y_t\}_{t=-\infty}^{\infty}$ exhibits no time dependence.

To verify that the process is a covariance stationary process we must show that the mean, variance and autocovariances are time invariant.

$$E[Y_t] = \mu + E[\varepsilon_t] + \theta E[\varepsilon_{t-1}] = \mu,$$

because $E[\varepsilon_t] = E[\varepsilon_{t-1}] = 0$.

For the variance, we have

$$\begin{aligned}
\mathrm{var}(Y_t) &= \sigma^2 = E[(Y_t - \mu)^2] = E[(\varepsilon_t + \theta\varepsilon_{t-1})^2] \\
&= E[\varepsilon_t^2] + 2\theta E[\varepsilon_t\varepsilon_{t-1}] + \theta^2 E[\varepsilon_{t-1}^2] \\
&= \sigma_\varepsilon^2 + 0 + \theta^2\sigma_\varepsilon^2 = \sigma_\varepsilon^2(1 + \theta^2).
\end{aligned}$$

The term $E[\varepsilon_t\varepsilon_{t-1}] = \mathrm{cov}(\varepsilon_t, \varepsilon_{t-1}) = 0$ because $\{\varepsilon_t\}_{t=-\infty}^{\infty}$ is an independent process.

For $\gamma_1 = \mathrm{cov}(Y_t, Y_{t-1})$, we have

$$\begin{aligned}
\mathrm{cov}(Y_t, Y_{t-1}) &= E[(Y_t - \mu)(Y_{t-1} - \mu)] \\
&= E[(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-1} + \theta\varepsilon_{t-2})] \\
&= E[\varepsilon_t\varepsilon_{t-1}] + \theta E[\varepsilon_t\varepsilon_{t-2}] \\
&\quad + \theta E[\varepsilon_{t-1}^2] + \theta^2 E[\varepsilon_{t-1}\varepsilon_{t-2}] \\
&= 0 + 0 + \theta\sigma_\varepsilon^2 + 0 = \theta\sigma_\varepsilon^2.
\end{aligned}$$

Note that the sign of γ_1 is the same as the sign of θ. For $\rho_1 = \mathrm{cor}(Y_t, Y_{t-1})$ we have

$$\rho_1 = \frac{\gamma_1}{\sigma^2} = \frac{\theta\sigma_\varepsilon^2}{\sigma_\varepsilon^2(1 + \theta^2)} = \frac{\theta}{(1 + \theta^2)}.$$

Example *Simulated values from MA(1) process*

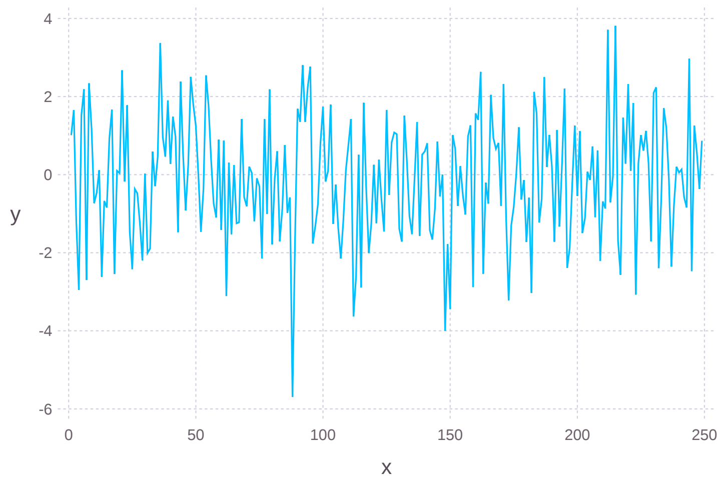
In [6]:
```
@manipulate for μ in -1:.1:1, θ in -1:.1:1, σ in 1:.1:2
    xs = collect(1:249)
    es = rand(Normal(0,σ), 250)
    ys = zeros(249)

    for t in 1:250
        if t > 1
            ys[t-1] = μ + es[t] + θ*es[t-1]
        end
    end

    plot(x=xs, y=ys, Geom.line)
end
```

μ 0.0

θ 0.0

σ 1.5

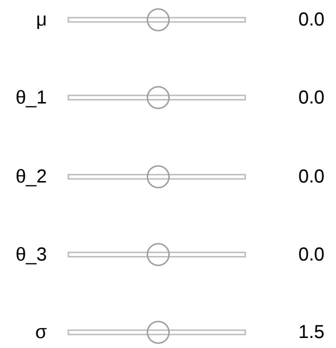
Out[6]:



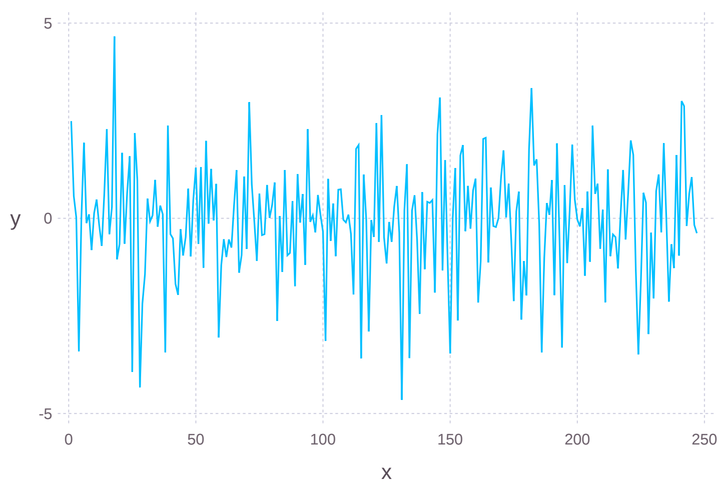
Example *Simulated values from MA(3) process*

In [7]:

```
@manipulate for μ in -1:.1:1, θ_1 in -1:.1:1, θ_2 in -1:.1:1, θ_3 in -1:.1:1, σ in 1:.1:2
    xs = collect(1:247)
    es = rand(Normal(0,σ), 250)
    ys = zeros(247)

    for t in 1:250
        if t > 3
            ys[t-3] = μ + es[t] + θ_1*es[t-1]+ θ_2*es[t-2] + θ_3*es[t-3]
        end
    end

    plot(x=xs, y=ys, Geom.line)
end
```



Out[7]:



1.3 Autoregressive Processes

1.3.1 AR(1) Model

Suppose you want to create a covariance stationary and ergodic stochastic process $\{Y_t\}_{t=-\infty}^{\infty}$ in which Y_t and Y_{t-1} are correlated and Y_t and Y_{t-2} are slightly less correlated, Y_t and Y_{t-3} are even less correlated and eventually Y_t and Y_{t-j} are uncorrelated for j large enough. That is, the time dependence in the process decays to zero as the random variables in the process get farther and farther apart. Such a process can be created using the first order autoregressive (AR(1)) model:

$$Y_t - \mu = \phi(Y_{t-1} - \mu) + \varepsilon_t, \quad -1 < \phi < 1, \quad \varepsilon_t \sim \; iid\, N(0, \sigma_\varepsilon^2)$$

It can be shown that the AR(1) model is covariance stationary and ergodic provided $-1 < \phi < 1$. We can also show that the AR(1) process has the following properties:

$$
\begin{aligned}
E[Y_t] &= \mu, \\
\mathrm{var}(Y_t) &= \sigma^2 = \sigma_\varepsilon^2/(1-\phi^2), \\
\mathrm{cov}(Y_t, Y_{t-1}) &= \gamma_1 = \sigma^2\phi, \\
\mathrm{cor}(Y_t, Y_{t-1}) &= \rho_1 = \gamma_1/\sigma^2 = \phi, \\
\mathrm{cov}(Y_t, Y_{t-j}) &= \gamma_j = \sigma^2\phi^j, \\
\mathrm{cor}(Y_t, Y_{t-j}) &= \rho_j = \gamma_j/\sigma^2 = \phi^j.
\end{aligned}
$$

Example *Simulated values from AR(1) process*

In [9]:

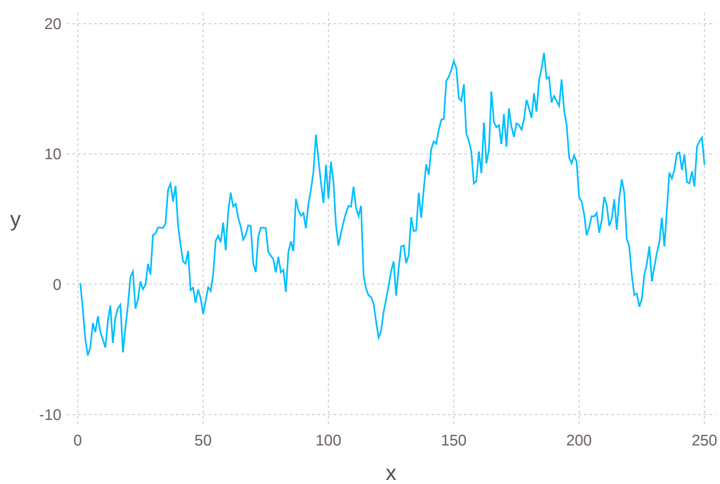
```
@manipulate for μ in -1:.1:1, ϕ in -1:.1:1, σ in 1:.1:2
    xs = collect(1:250)
    es = rand(Normal(0,σ), 250)
    ys = zeros(250)

    for t in 1:250
        if t <= 1
            ys[t] = μ - ϕ*μ + es[t]
        else
            ys[t] = μ - ϕ*μ ys[t-1] + es[t]
        end
    end

    plot(x=xs, y=ys, Geom.line)
end
```

μ -0.4

ϕ 0.8

σ 1.7

Out[9]:



Example *Simulated values from AR(3) process*

In [19]:

```
@manipulate for μ in -1:.1:1, ϕ_1 in -1:.1:1, ϕ_2 in -1:.1:1, ϕ_3 in -1:.1:1, σ in 1:.1:2
    xs = collect(1:250)
    es = rand(Normal(0,σ), 250)
    ys = zeros(250)

    for t in 4:250
        ys[t] = μ + ϕ_1*ys[t] + ϕ_2*ys[t-1] + ϕ_3*ys[t-3]+es[t-2]
    end

    plot(x=xs, y=ys, Geom.line)
end
```

μ 0.0

ϕ_1 0.5

ϕ_2 0.5

ϕ_3 -0.7

σ 1.7

Out[19]: