WDS'06 Proceedings of Contributed Papers, Part I, 82–87, 2006. *ISBN 80-86732-84-3 © MATFYZPRESS*

Stochastic Epidemic Models

J. Staněk

Charles University, Faculty of Mathematics and Physics, Prague, Czech Republic.

Abstract. The present work concerns models of epidemic which are described by differential and stochastic differential equations. The theorems on the existence, uniqueness and properties of their solutions are stated.

Introduction

In nature, there are many diseases which can cause epidemics. Therefore it is useful to model the behavior of these diseases in order to control their progression. In this paper, the research is aimed especially at some models of epidemic described by stochastic differential equations.

In the first part, the Kermack-McKendrick deterministic model is stated. The second part describes its stochastic version. The third part concerns a model for multiple diseases. In all these parts, the main aim is to find conditions for the existence and uniqueness of the solution to the describing equations. Furthermore, properties of these solutions are described.

Kermack-McKendrick deterministic model

The Kermack-McKendrick model is a classical deterministic epidemic model. This model assumes a fixed size of population n. That is divided into three subpopulation: *susceptibles* (the individuals who can be infected), *infectives* (the infected individuals who are able to spread the disease) and *removals* (the individuals restore to health not able further to spread the infection or get themselves to be infected again). Denote $x(t)$ the size of the susceptibles subpopulation in time t, $y(t)$ the size of the infectives subpopulation at time t and $z(t)$ the size of the removals subpopulation at time t. Hence $x(t) + y(t) + z(t) = n$. The model is given by the following three dimensional differential equation:

$$\begin{aligned} dx(t) &= -\beta x(t)y(t)dt, \quad x(0) = x_0 > 0, \\ dy(t) &= \beta x(t)y(t)dt - \gamma y(t)dt, \quad y(0) = y_0 > 0, \\ dz(t) &= \gamma y(t)dt, \quad z(0) = 0, \end{aligned} \tag{R1}$$

where the intensity $\beta > 0$ is the pairwise rate of infection and $\gamma > 0$ is the removal rate of infectives.

For this model the following results can be proved.

Theorem 1 *(R1) has a unique solution* $(x, y, z) \in \mathcal{C}^1(\mathbf{R}^+, \mathbf{R}^3)$ *that is positive on* $(0, \infty)$ *and such that*

$$\begin{aligned} x(t) &= x_0 \exp\left\{-\frac{\beta}{\gamma}z(t)\right\}, \\ y(t) &= n - z(t) - x(t). \end{aligned}$$

The number of susceptibles $x(t)$ *is a nonincreasing function on* $(0, \infty)$ *and* $z(t)$ *in an increasing function on* $(0, \infty)$*. The limits* x_∞, y_∞ *and* z_∞ *exist, and* $y_\infty = 0$*.*

Proof Obviously, $x(t)$ and $y(t)$ are solutions to (R1). Then using Theorem 1 from [*Štěpán and Hlubinka, 2006*] we get the existence and uniqueness of solution (R1) and the relations for the limits.

Stochastic Models

Stochastic version of Kermack-McKendrick model

Consider a model of epidemic with unstable size of the population N_t for which the inequality $0 < a \leq N_t \leq b < \infty$ holds for all $t > 0$, and the population being divided into three subpopulations which change their respective sizes in the running time: X_t is the size of susceptibles, Y_t is the size of infectives and Z_t is the size of removals. This model is given by the following three dimensional stochastic differential equation

$$\begin{aligned} dX_t &= -\beta(X_t,Y_t,Z_t)X_tY_tdt + X_t\sigma(X_t+Y_t+Z_t)dW_t, && X_0 = x_o > 0,\\ dY_t &= \beta(X_t,Y_t,Z_t)X_tY_tdt - \gamma Y_tdt + Y_t\sigma(X_t+Y_t+Z_t)dW_t. && Y_0 = y_0 > 0, \qquad \text{(R2)}\\ dZ_t &= \gamma Y_tdt + Z_t\sigma(X_t+Y_t+Z_t)dW_t, && Z_0 = 0, \end{aligned}$$

where the intensities $\beta(x,y,z) > 0$, $\gamma > 0$ are described above, $\sigma(.) \geq 0$ is a diffusion coefficient and W_t is a standard Brownian motion. Note that the size N_t is a solution to the Engelbert-Schmidt stochastic differential equation

$$dN_t = N_t\sigma(N_t)dW_t, \quad N_0 = n_0 = x_0 + y_0.$$

The following theorem holds.

Theorem 2 *Assume $\beta(x,y,z)$ and $\sigma(n)$ bounded. Then (X,Y,Z) is a solution to (R2) if and only if*

$$\begin{aligned} N_t &= n_0 \exp\left\{\int_0^t \sigma(N_u)dW_u - \frac{1}{2}\int_0^t \sigma^2(N_u)du\right\},\\ X_t &= \frac{x_0}{n_0}\exp\left\{-\int_0^t \beta(X_u,Y_u,Z_u)Y_udu\right\}\cdot N_t,\\ Y_t &= \frac{y_0}{n_0}\exp\left\{\int_0^t \beta(X_u,Y_u,Z_u)X_udu - \gamma t\right\}\cdot N_t,\\ Z_t &= \gamma\int_0^t \frac{Y_u}{N_u}du\cdot N_t \end{aligned}$$

hold for all $t \geq 0$ almost surely. Especially, the processes X, Y and N_t are positive on $\mathbf{R}^+$ and Z_t is a positive process on $(0,\infty)$.

Some basic properties of the solution to (R2) are described by following theorem.

Theorem 3 *Assume β nonnegative and bounded, that supp(σ) $\subset [a,b]$ and consider an arbitrary solution (X,Y,Z) to (R2). Then the following statements hold:*

(i) The size of population N_t is a bounded martingale such that $a \leq N \leq b$ holds and such that the limit $N_\infty := \lim_{t\to\infty} N_t$ exists.

(ii) The number of the susceptibles X_t is a supermartingale and the number of removals Z_t is a submartingale.

(iii) All limits X_∞, Y_∞ and Z_∞ exist, $Y_\infty = 0$, and $X_\infty > 0$ if and only if $N_\infty > 0$, hence assuming $a > 0$ we get X_∞ as a positive random variable.

Both theorems are proved in [*Štěpán and Hlubinka*, 2006].

Remark The existence and uniqueness of solution (R2) follows from the more general model described below by (R3).

Remark In the model described by (R2), the number of newly infected people at time t depends only on the size of X, Y, Z in time t. But in fact, the number of newly infected people at time t can depend on the size of X, Y, Z for the whole time period $[0,t]$. That is the reason, why we should be interested in choosing more complex intensities β which will be shown below.

Consider a generalization of the model given by (R2), described by the next three dimensional stochastic differential equation:

$$\begin{aligned} dX_t &= -\beta(X_.,Y_.,Z_.,t)X_tY_tdt + X_t\sigma(X_t+Y_t+Z_t)dW_t, && X_0 = x_o > 0,\\ dY_t &= \beta(X_.,Y_.,Z_.,t)X_tY_tdt - \gamma Y_tdt + Y_t\sigma(X_t+Y_t+Z_t)dW_t, && Y_0 = y_0 > 0, \qquad \text{(R3)}\\ dZ_t &= \gamma Y_tdt + Z_t\sigma(X_t+Y_t+Z_t)dW_t, && Z_0 = 0, \end{aligned}$$

where $\beta(x_., y_., z_., t) : \mathcal{C}(\mathbf{R}^+, \mathbf{R}^3) \times \mathbf{R}^+ \to \mathbf{R}^+$ is a progressive path functional. This means that $\beta(X_., Y_., Z_., .)$ is an $\mathcal{F}_t$-progressive process on $\mathcal{C}(\mathbf{R}^+, \mathbf{R}^3)$, where $\mathcal{F}_t$ is the minimal right continuous filtration built up on the top of the canonical filtration $\mathcal{H}_t$ on $\mathcal{C}(\mathbf{R}^+, \mathbf{R}^3)$, and $\sigma(.)$ and γ are as in the preceding model. Further assume that $0 < a \le N_t \le b < \infty$ holds for all $t \ge 0$.

Denote $\triangle_{ab} := \{(x_., y_., z_.) \in \mathcal{C}(\mathbf{R}^+, \mathbf{R}^3) : a \le x_t + y_t + z_t \le b \text{ a } x_t, y_t, z_t \ge 0 \text{ for all } t \ge 0\}$ and assume that $\beta(.)$ and $\sigma(.)$ satisfy the following conditions:

(i) $\beta(x_., y_., z_., t)$ is a progressive path functional from $\mathcal{C}(\mathbf{R}^+, \mathbf{R}^3) \times \mathbf{R}^+$ to $\mathbf{R}^+$,

(ii) $\beta(.)$ is bounded and locally Lipschitz on $\triangle_{ab}$, i.e. for all $N \in \mathbf{N}$ there exist constants K_N and K such that

$$|\beta(l_., s) - \beta(\tilde{l}_., s)| \le K_N ||l - \tilde{l}||_s^*,$$
$$|\beta(l_., s)| \le K$$

holds for all $l_., \tilde{l}_. \in \triangle_{ab}$, for which $||l||_s^* \vee ||\tilde{l}||_s^* \le N$ and $0 \le s \le N$,where $l_. = (x_., y_., z_.)$ is a continuous function from $\mathbf{R}^+$ to $\mathbf{R}^3$, $|\cdot|$ is the Euclidean norm on $\mathbf{R}^3$ and $||f||_s^* \equiv \sup\{|f(t)| : t \le s\}$,

(iii) $\sigma(.)$ is a bounded measurable function on $\mathbf{R}^+$ supported by a compact$[a, b]$,

(iv) $\sigma(.)$ is locally Lipschitz on $\mathbf{R}^+$.

Now we can introduce the theorem on the existence of the unique strong solution of (R3).

Theorem 4 *Let $\sigma(.)$ and $\beta(.)$ satisfy the conditions (i)-(iv). Then the equation (R3) has a unique strong solution and an arbitrary solution $L_t = (X_t, Y_t, Z_t)$ is a nonnegative process on $(0, \infty)$.*

This theorem is proved in [*Staněk*, 2006].

As mentioned above, the number of the newly infected people at time t can depend on the size of X, Y, Z for some time period before time t (incubational time period). This model allows to describe situations, when the disease has a long incubational time period, during which the disease can burn up.

Example Consider a disease which has the incubation time period $\widetilde{t}$ and that is such that during this time the disease can burn up with the same intensity. Then a possible choice of the coefficient $\beta(.)$ is

$$\beta(x_., y_., z_., t) = C \int_{t-\tilde{t}}^{t} \frac{x_u y_u}{(x_u + y_u + z_u)^2} du.$$

Choosing such a $\beta(.)$, the number of newly infected people at time t depends on the sizes X_t, Y_t and on the size of the rates $\frac{X}{N}, \frac{Y}{N}$ for the whole time period $(t - \widetilde{t}, t)$.

Another possible choice of $\beta(.)$ is

$$\beta(x_., y_., z_., t) = C \int_{t-\tilde{t}}^{t} \frac{y_u}{x_u + y_u + z_u} du,$$
$$\beta(x_., y_., z_., t) = C \int_{t-\tilde{t}}^{t} y_u du.$$

Model with multiple pathogens

Consider again a model of epidemic with unstable size of the population N_t. This model will be used for disease with multiple pathogens. Suppose the population being divided into $n+1$ subpopulations: X_t is the size of susceptibles and $Y_t^j,\ j = 1, ..., n$, is the size of population which is infected by the pathogen strain $j,\ j = 1, ..., n$. This model assumes that nobody is infected by two or more pathogen strains and somebody infected by some pathogen strain is immune for the other pathogen strains. This situation is called the cross immunity. We consider the end of epidemic as the stopping time τ_f when some of processes X_t, Y_t^j enters to zero or to N^{bound}, where $N^{bound} > N_0$ is a chosen (constant) upper bounder for the size of population. The first condition means that there are no people who can be infected or that some disease has finished, so we stop the model. The second condition means that the size of population overgrows some boundary. We use these conditions to show the existence and uniqueness of a solution.

The model is given by the following $n+1$ stochastic differential equation

$$
\begin{aligned}
dX_t &= X_t\left(b - d(N_t) - \sum_{k=1}^{n}\frac{\beta_k Y_t^k}{N_t}\right)dt + \sum_{k=1}^{n} b_k Y_t^k dt + \sum_{k=1}^{n+1} B_{1.k}(L_t)dW_t^k, \\
dY_t^j &= Y_t^j\left(b - b_j - d(N_t) - \alpha_j + \frac{\beta_j X_t}{N_t}\right)dt + \sum_{k=1}^{n+1} B_{j+1.k}(L_t)dW_t^k, \quad j = 1,\ldots,n,
\end{aligned}
\tag{R4}
$$

for $t \in (0, \tau_f)$, where b is the per capita birth rate, the function $d(.)$ is the per capita death rate, and for $j = 1,\ldots,n$, b_j is the birth rate of health offspring to parents infected by j-th disease, α_j is the j-th disease-related, per capita death rate and β_j is the transmission rate for j-th disease. Next, W_t^k, $k = 1,...,n+1$, are independent Wiener processes, $L_t = (X_t, Y_t^1,...,Y_t^n)$ and $B(L_t) = \sqrt{C(L_t)}$, where $C(L_t)$ is a $(n+1)\times(n+1)$-matrix, which is symmetric, positive definite and is defined by

$$
C(L_t) = \begin{pmatrix} 0 & \sigma_{1.2}(L_t) & \cdots & \sigma_{1.n+1}(L_t) \\ \sigma_{2.1}(L_t) & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n+1.1}(L_t) & 0 & \cdots & 0 \end{pmatrix} + \operatorname{diag}(\sigma_{11}(L_t), \sigma_{22}(L_t), \ldots, \sigma_{n+1,n+1}(L_t)),
$$

where

$$
\begin{aligned}
\sigma_{1.1}(L_t) &= X\left(b + d(N) + \sum_{k=1}^{n}\frac{\beta_k Y^k}{N}\right) + \sum_{k=1}^{n} b_k Y^k, \\
\sigma_{j+1.j+1}(L_t) &= Y^j\left(b - b_j + d(N) + \alpha_j + \beta_j\frac{X}{N}\right), \quad j = 1,...,n, \\
\sigma_{1.j+1}(L_t) &= -\beta_j\frac{XY^j}{N} = \sigma_{j+1.1}, \quad j = 1,...,n.
\end{aligned}
$$

Matrix $C(L_t)$ is the covariance matrix for the change in the population sizes. For $j,k = 2,...,n$, the coefficients $\sigma_{1,k+1}(L_t)$ describe the interaction between X_t and Y_t^k and $\sigma_{j,k}(L_t) = 0$ for $j \neq k$ because of the cross immunity. It can be proved, that $C(L_t)$ is strictly positive definite if and only if $X, Y^j > 0$ and $b, d(N), b - b_j, \alpha_j > 0$ holds. Under these conditions $C(L_t)$ is positive definite, and therefore, it has a unique positive definite square root matrix $B(L_t)$. For this model, it is possible to prove the following theorem on the existence and uniqueness of a solution.

Theorem 5 *Let $X(0), Y^k(0) > 0$ for $k = 1,\ldots,n$, $d(.)$ is a bounded local Lipschitz function and $b, d(N), b - b_j, \alpha_j > 0$. Then there exists a stopping time τ, such that the equation (R4) has a unique solution for $t \in (0,\tau)$. If $L_t = (X_t, Y_t^1,...,Y_t^n)$ is a solution of (R4), a possible choice of τ is $\tau = \tau_f$, where τ_f is described above.*

Proof The equation (R4) can be rewritten as

$$
dL_t = a(L_t)dt + B(L_t)dW_t, \tag{R5}
$$

where $W_t = (W_t^1,...,W_t^{n+1})$, $a(L_t) = (a^1(L_t),...,a^{n+1}(L_t))$ and

$$
\begin{aligned}
a^1(L_t) &= X_t\left(b - d(N_t) - \sum_{k=1}^{n}\frac{\beta_k Y_t^k}{N_t}\right) + \sum_{k=1}^{n} b_k Y_t^k, \\
a^j(L_t) &= Y_t^j\left(b - b_j - d(N_t) - \alpha_j + \frac{\beta_j X_t}{N_t}\right), \quad j = 2,..,n+1.
\end{aligned}
$$

Let $1 > \epsilon > 0$. Denote $\widehat{l} = ((l_1 \vee \epsilon)\wedge N_b, ..., (l_{n+1} \vee \epsilon)\wedge N_b)$ for some vector $l = (l_1,\ldots,l_{n+1})$. Define a map $\widehat{C} : \mathbf{R}_+^{n+1} \to \mathbf{M}^{n+1\times n+1}$, so that $\widehat{C}(l) = C(\widehat{l})$. Then $\widehat{C}(l)$ is everywhere strictly positive definite, and from theorem 12.12 in [*Rogers and Williams, 1994*], p.134, the map $\widehat{B} := \widehat{C}^{\frac{1}{2}}$ is local Lipschitz. Thus, for all N there exists K_N such that

$$
|\widehat{B}(l) - \widehat{B}(\widetilde{l})| \le K_N |l - \widetilde{l}| \tag{1}
$$

holds for all $l, \widetilde{l}$ satisfying $|l|, |\widetilde{l}| \leq N$, where $|\widehat{B}(l)| \equiv \sqrt{\mathrm{trace}(\widehat{B}(l)\widehat{B}(l)^T)}$ and $|l|$ is Euclidean norm of l.

Because for all $j = 1, ..., n+1$, $a^j(l)$ are local Lipschitz, we get that $a(l)$ is local Lipschitz, and so for all N there exists K_N such that

$$|a(l) - a(\widetilde{l})| \leq K_N |l - \widetilde{l}| \tag{2}$$

holds for all $l, \widetilde{l}$ satisfying $|l|, |\widetilde{l}| \leq N$.

Since $d(.) \leq K_d$ for some $K_d < \infty$ then

$$|a(l)| \leq \sqrt{\left(\sum_{k=1}^{n}(\beta_k + b_k) + b + K_d\right)^2 + \sum_{k=1}^{n}(b - b_k + K_d + \alpha + \beta)^2 |l|} \leq K|l| \tag{3}$$

holds for some $K < \infty$.

Now, we show that $b_{i.j}$ are bounded for all $i, j = 1, \ldots, n+1$. Denote $\mathbf{1} = (1, ..., 1)$ the $n+1$ dimensional vector. Then $\widehat{B}(\mathbf{1}) = B(\mathbf{1})$, $\sigma_{i.j}(\mathbf{1})$ are bounded constants for all $i, j = 1, ..., n+1$ and $\sigma_{i.i} > 0$ for all $i = 1, ..., n+1$. From the Cholesky decomposition we have:

$$\widehat{b}_{i,i}(\mathbf{1}) = b_{i.i}(\mathbf{1}) = \sqrt{\sigma_{i.i}(\mathbf{1}) - \sum_{k=1}^{i-1} b_{i.k}^2(\mathbf{1})},$$

$$\widehat{b}_{k,i}(\mathbf{1}) = b_{j.i}(\mathbf{1}) = (\sigma_{j.i}(\mathbf{1}) - \sum_{k=1}^{i-1} b_{j.k}(\mathbf{1}) b_{i.k}(\mathbf{1}))/b_{i.i}(\mathbf{1}).$$

Because $\widehat{b}_{1,1}(\mathbf{1})$ is a finite positive constant then for all $i = 2, ..., n+1$, $\widehat{b}_{i,1}(\mathbf{1})$ are a bounded constants, so for the first column, the boundary is shown. Suppose (for induction) that for $k = 1, ..., n+1$ and $l = 1, ..., i-1$, $b_{k.l}(\mathbf{1})$ are bounded and $b_{l.l}(\mathbf{1})$ are positive bounded constants. Let $b_{i.i}(\mathbf{1}) = 0$. Then $b_{i+1.i}(\mathbf{1}) = \infty$, and therefore

$$\sigma_{i+1,i+1}(\mathbf{1}) = \sum_{j=1}^{i+1} b_{i+1.j}^2(\mathbf{1}) \geq b_{i+1.i}^2(\mathbf{1}) = \infty.$$

But this is disagreement, because $\sigma_{i,j}$ are bounded for all $i, j = 1, ..., n+1$, therefore $b_{i.i}(\mathbf{1}) \neq 0$. Again from the Cholesky decomposition, we get that $b_{j.i}(\mathbf{1})$ are bounded for all $j = i+1, .., n+1$, so using the induction we get that $b_{i,j}$ are bounded for all $i, j = 1, ..., n+1$, and therefore, there exist a constant $C_B < \infty$ satisfying

$$|\widehat{B}(\mathbf{1})| \leq C_B < \infty. \tag{4}$$

Using (1) and (4) we get that for all N, there exists K_N such that

$$|\widehat{B}(l)| \leq |\widehat{B}(l) - \widehat{B}(\mathbf{1}) + \widehat{B}(\mathbf{1})| \leq K_N |l - \mathbf{1}| + C_B \leq C_N(1 + |l|) \tag{5}$$

holds for all l satisfying $|l| \leq N$. Since $\widehat{B}(l) = \widehat{B}(\widetilde{l})$, $|\widetilde{l}| \leq \sqrt{n} N_b$ and using (5) we get that

$$|\widehat{B}(l)| \leq K(1 + |l|) \tag{6}$$

holds for some $K < \infty$. Using the properties (1)-(3) and (6), then due to the theorem 12.1 in [*Rogers and Williams, 1994*], p.132, the equation

$$dL_t = a(L_t)dt + \widehat{B}(L_t)dW_t \tag{R6}$$

has a unique solution. Note that the stoping time τ_ϵ is the time of the first output of the process L_t from the interval $(\epsilon, N_b)^{[n+1]}$. Then, because

$$L_{t\wedge\tau_\epsilon} = L_0 + \int_0^{t\wedge\tau_\epsilon} a(L_s)ds + \int_0^{t\wedge\tau_\epsilon} \widehat{B}(L_s)dW_s = L_0 + \int_0^{t\wedge\tau_\epsilon} a(L_s)ds + \int_0^{t\wedge\tau_\epsilon} B(L_s)dW_s,$$

the equation (R4) has a unique solution in the period $[0, \tau_\epsilon]$. Finally, for $\epsilon \to 0$ we get the unique solution to (R4) in the time period $[0, \tau_f)$.

Conclusion

In this work, some models of epidemic described by differential and stochastic differential equations were stated. The Kermack-McKendrick model and his stochastic version is well established. In the case of multiple diseases, there stay two questions. The first is how to extend the model for a longer time interval. The second question is describing of the properties of the solution.

Acknowledgments. This work was supported by grant the 201/05/H007. The author also thanks to Prof. Josef Štěpán for the professional help.

References

Allen L.J.S. and Kirupaharan N., Asymptotic Dynamics of Deterministic and Stochastic Epidemic Models with Multiple Pathogens, *International Journal of Numerical Analysis and Modeling*, 3(2), 329-344, 2005.

Busenberg S. and Kenneth C., Vertically Transmitted Diseases - Models and Dynamics, *Springer Verlag*, Berlin-Heidelberg-New York, 1993.

Daley D.J. and Gani J., Epidemic Modelling: An Introduction, *Cambridge University Press*, Cambridge, 1999.

Rogers L.C.G. and Williams D., Diffusions, Markov Processes and Martingales, Volume 2 - Itô Calculus, *John Wiley and Sons*, Chichester-New York-Brisbane-Toronto-Singapore, 1994.

Staněk J., Stochastická verze klasického modelu vývoje epidemie [in Czech], in: Proceedings of *Robust 2006*, Praha, 2006.

Štěpán J. and Hlubinka D., Kermack-McKendrick Epidemic Model Revisited, *Kybernetika*, 42, 2006, to appear.