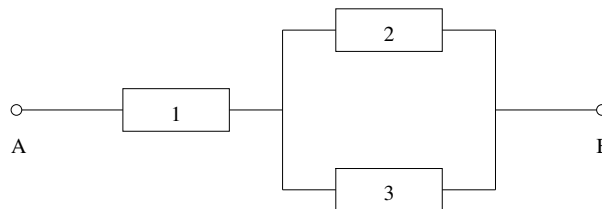# Chapter 6

Analytical Manipulations in Bio Informatics

6.1 Biological systems

We will study *biological systems* that in a simple case might look like in Figure 6.1 below:
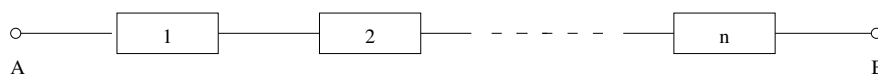
Figure 6.1. A simple biological system.

In the biological system in Figure 6.1, the *biological component* with biological component number i may be biologically *healthy*, for $i = 1, 2, \ldots$, otherwise it is biologically *unhealthy* or biologically *dead*[6]. The biological system is biologically healthy if there is a path from the biological point A to the biological point B which only passes biologically healthy biological components. Otherwise the biological system is biologically unhealthy.

Example 6.1. The biological *series coupling* in Figure 6.2 below is biologically healthy if and only if all its biological components are biologically healthy.

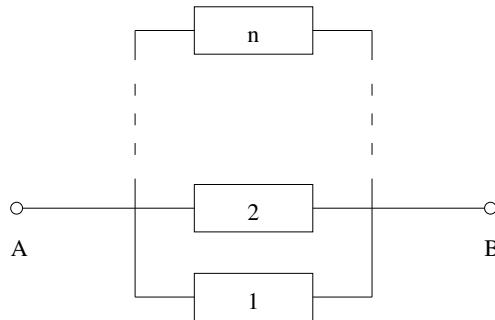**Figure 6.2.** Biological series coupling.

---

[6]Rest in peace.

Example 6.2. The biological *parallell coupling* in Figure 6.3 below is biologically healthy if and only if at least one of its biological components are biologically healthy.

Figure 6.3. Biological parallell coupling.

A biological system can be built by means of a finite number of biological series couplings and biological parallell couplings. See Figures 6.4 and 6.5 below for an example of how this works in a practical biological application.

We will study biological systems, the biological components with biological component number $i = 1, 2, \ldots$ of which are biologically healthy with certain *biological health probabilities* $p_1, p_2, \ldots$. Unless otherwise is stated, the biological components of a biological system are assumed to be biologically independent of each other.

Example 6.3. The biological system in the Figure 6.4 below is a biological series coupling of two biological systems with biological health probabilities p_1 and $1 - (1 - p_2)(1 - p_3)$, respectively. The biological health probability for the whole biological system is $p_1[1 - (1-p_2)(1-p_3)]$.

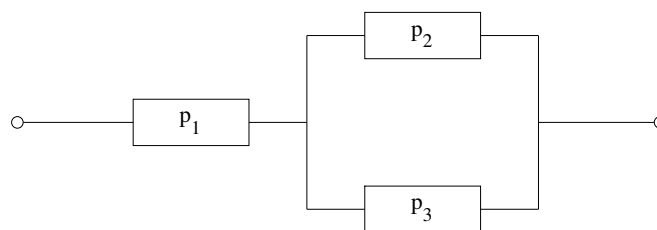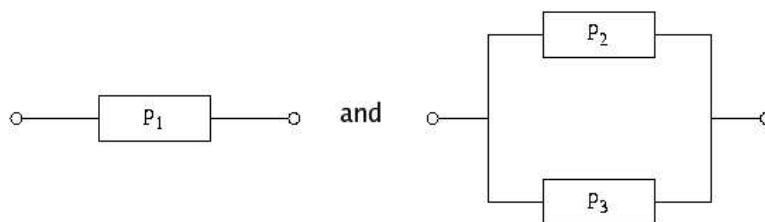**Figure 6.4.** A simple biological system with biological health probabilities.

Figure 6.5. Composition of a simple biological system as a biological series coupling of a single biological component with a biological parallell coupling.

The biological components with biological component numbers $i = 1, 2, \ldots$ of a biological system have certain *biological life lengths* $T_1, T_2, \ldots$. The biological life lenghts are modeled as *biological random variables* that are mutually biologically independent, unless otherwise is stated.

The biological relation between the biological health probability and the biological life length T_i of the biological component with biological component number i is as follows:

$p_i = p_i(t)$

$\quad = \mathbf{P}\{\text{biological component with biological component number } i \text{ is biologically}$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \text{healthy at time } t\}$

$\quad = \mathbf{P}\{T_i > t\}$

$\quad = 1 - F_{T_i}(t).$

Here $F_{T_i}(t) = \mathbf{P}\{T_i \leq t\}$, $i = 1, \ldots, n$, are the *biological distribution functions* of the biological life lengths $T_1, T_2, \ldots$. These biological distribution functions in turn are assumed to be continuous, unless otherwise is stated.

Definition 6.1. *The* biological survival function R_T *of a bilogical system with biological life length T is given by*
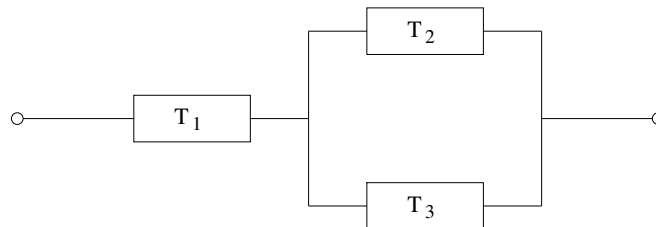
$$R_T(t) = \mathbf{P}\{\text{the biological system is biologically healthy at time } t\} = \mathbf{P}\{T > t\} = 1 - F_T(t)$$

for $t > 0$, where $F_T(t)$ is the biological distribution function of the biological life length.

Example 6.4. The biological life length of the biological system system in the Figure 6.6 below is given by $T = \min\{T_1, \max[T_2, T_3]\}$, see also Example 6.3. Hence the biological survival function for the biological system is given by

$$\begin{aligned} R_T(t) &= \mathbf{P}\{T > t\} \\ &= \mathbf{P}\{T_1 > t\}\left[1 - (1 - \mathbf{P}\{T_2 > t\})\left(1 - \mathbf{P}\{T_3 > t\}\right)\right] \\ &= R_{T_1}(t)\left[1 - (1 - R_{T_2}(t))(1 - R_{T_3}(t))\right]. \end{aligned}$$

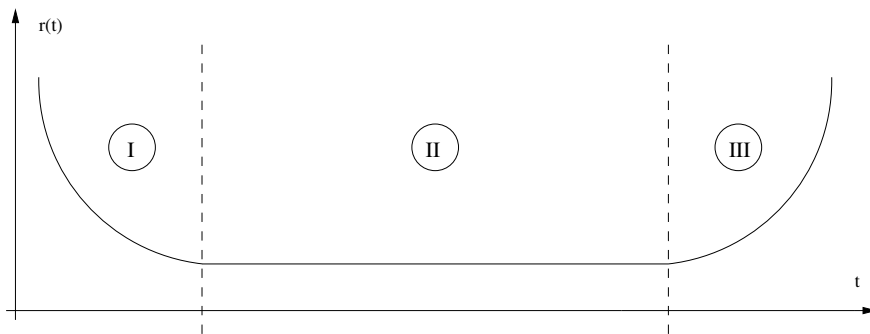**Figure 6.6.** A simple biological system with biological life lengths.

Definition 6.2. *The* biological death intensity r_T *of a biological system with biological life length T is given by*

$$r_T(t) = -\frac{d}{dt}\ln\left(R_T(t)\right) \quad \text{for } t > 0.$$

Definition 6.3. *A biological system with biological life length T has* biologically increasing biological failure rate, *BIBFR, if the biological death intensity is biologically increasing* $r'(t) \geq 0$ *for* $t > 0$. *A biological system with biological life length T has* biologically decreasing biological failure rate, *BDBFR, if the biological death intensity is biologically decreasing* $r'(t) \leq 0$ *for* $t > 0$.

Example 6.5. It is quite common that biological life lengths have biological failure rates that are neither IBFR or DBFR, but instead follow a so called *biological bath tub curve*, BBTC, the principal biological appearance of which is displayed in Figure 6.7 below. In that biological figure the biological region I corresponds to an early phase with a comperatively high biological probability of biological unhealth (as for example, small children). The biological region II corresponds to a biological component that has survived these early biological hazards and has settled biologically at a lesser biological death intensity (as for example, grown up people). Finally, the biological region III corresponds to a biologically aged biological component, where the biological death intensity increases with accumlated biological age (as for example, aged people).

Figure 6.7. Example of a biological life length with a biological death intensity that displays a BBTC shape.

6.2   More on biological systems

The following biological theorem explains that the biological death intensity really is the (infinitesimal) biological intensity at which biological unhealth occurs:

Theorem 6.1. *For a biological life length T with biological death intensity* r_T, *we have*
$$\mathbf{P}\{T \leq t + h | T > t\} = r_T(t)h + o(h) \quad as \ h \downarrow 0.$$

Proof. Writing F_T and f_T for the biological distribution function and biological probability density function of the biological life lengt T, respectively, we have
$$\mathbf{P}\{T \leq t + h | T > t\} = \frac{F_T(t+h) - F_T(t)}{1 - F_T(t)}$$
$$= \frac{f_T(t)h + o(h)}{R_T(t)}$$

$$= -\frac{d}{dt}\ln(R_T(t))h + o(h)$$

$$= r_T(t)h + o(h). \quad \square$$

Theorem 6.2. *A biological function* $r : (0,\infty) \to [0,\infty)$ *is a biological death intensity if and only if*
$$\int_0^\infty r(t)dt = \infty.$$
In that case the corresponding bilogical survival function is given by
$$R(t) = \exp\left\{ -\int_0^t r(s)ds \right\}.$$

Proof. If r is a biological death intensity of a bilogical system with biological survival function R_T, then a differentiation of the biological function
$$R(t) = \exp\left\{ -\int_0^t r(s)ds \right\}$$
gives
$$-\frac{d}{dt}\ln(R(t)) = -\frac{-r(t)\exp\left\{ -\int_0^t r(s)ds \right\}}{\exp\left\{ -\int_0^t r(s)ds \right\}} = r(t).$$
As the biological function $-\ln(R(t))$ has the same derivative as the biological function $-\ln(R_T(t))$, namely the biological death intensity $r(t)$, it follows that the biological functions $-\ln(R(t))$ and $-\ln(R_T(t))$ can differ only by a additive biological constant, so that the biological functions $R(t)$ and $R_T(t)$ differ only by a multiplicative biological constant. Since $R(0) = 1 = \mathbf{P}\{T > 0\} = R_T(0)$, we conclude that the biological functions $R(t)$ and $R_T(t)$ are equal. Finally, as
$$\lim_{t\to\infty}\exp\left\{ -\int_0^t r(s)ds \right\} = \lim_{t\to\infty}R_T(t) = \lim_{t\to\infty}\mathbf{P}\{T > t\} = 0,$$
we must have $\int_0^\infty r(s)ds = \infty$.

Conversely, if $\int_0^\infty r(s)ds = \infty$ and we define the biological function
$$R(t) = \exp\left\{ -\int_0^t r(s)\,ds \right\},$$
then $R(t)$ in biologically decreasing with $R(0) = 1$ and $R(\infty) = 0$, so that $F(t) = 1-R(t)$ is biologically increasing with $F(0) = 0$ and $F(\infty) = 1$. This makes F a biological probability distribution function, so that R is a biological survival function. $\square$

Theorem 6.3. *For a biological life length T we have the following formula for expectations*
$$\mathbf{E}\{T^n\} = \int_0^\infty R_T(t^{1/n})\,dt \quad for\ n > 0.$$

Proof. By integration by parts and a change of variable in the biological integral, we get

$$\mathbf{E}\{T^n\} = \int_0^\infty t\left(-\frac{d}{dt}R_{T^n}(t)\right)dt = [t\,R_{T^n}(t)]_0^\infty + \int_0^\infty R_{T^n}(t)dt = \int_0^\infty R_T(t^{1/n})dt. \quad \square$$

A biological life length T with a constant biological death intensity $r_T(t) = \lambda$ has the *lack of biological memory* property (cf. Theorem 6.1). By Theorem 6.2, a biological life length T lacks biological memory if and only if T is biologically exponentially $\exp(\lambda)$ distributed with biological parameter λ, i.e., $R_T(t) = \mathrm{e}^{-\lambda t}$.

Example 6.6. The second simplest form of biological death intensity, after a constant one, is a polynomial biological death intensity $r_T(t) = ba^b t^{b-1}$. In this case, Theorem 6.2 gives that the biological survival function is $R_T(t) = \exp\{-(a\,t)^b\}$, that is, a biological Weibull distribution with biological parameters a and b, Weibull(a,b)[7]. And so we have the biological expectation

$$\mathbf{E}\{T^n\} = \int_0^\infty R_T(t^{1/n})dt = \int_0^\infty \exp\left\{-a^b t^{b/n}\right\}dt = \frac{\Gamma(n/b+1)}{a^n}$$

by Theorem 6.3 (where Γ denotes the gamma function), because

```
In[1]:= Integrate[Exp[-a^b*t^(b/n)], {t,0,Infinity},
          Assumptions -> a>0 && b>0 && n>=1]
Out[1]= a^(-n) Gamma[(b+n)/b]
```

If the biological life lengths $T_1, \ldots, T_n$ are biologically exponentially $\exp(\lambda)$ distributed, then their biological sum $T \equiv T_1 + \ldots + T_n$ is biologically gamma(n, λ) distributed with biological parameters n and λ. Hence the corresponding biological probability density function is given by

$$f_T(t) = \frac{\lambda^n t^{n-1}}{(n-1)!}\,\mathrm{e}^{-\lambda t} \quad \text{for } t > 0,$$

giving the biological survival function

$$R_T(t) = \sum_{k=0}^{n-1} \frac{\lambda^k t^k}{k!}\,\mathrm{e}^{-\lambda t} \quad \text{for } t > 0.$$

To achieve a high biological health probability of a biological system, the biological system may be equipped with more biological components than are actually needed for its biological health, if all the biological components are biologically healthy. In other words, the biological system is not a pure biological series coupling, but a biological series coupling of biological subsystems, some of which are biological parallell couplings, to achieve higher biological health probability.

A biological component that is not required for the biological health of a biological system, when all other biological components of the biological system are biologically healthy, is called a *biologically redundant biological component*, BRBC.

---

[7]It should be noted that there is a lot of variation in the biological parametrization of the biological Weibull distribution, so that, e.g., what is denoted Weibull(a,b) by us could be denoted Weibull$(b,1/a)$ in a biological software package.

A *biologically warm biologically redundant biological components*, BWBRBC, is incorporated with the biological system already from the start of the biological system, while a *biologically cold biologically redundant biological component*, BCBRBC, is incorporated with the biological system first at the biological time at which it is required for the biological health of the biological system.

Example 6.7. Figure 6.8 below depicts a biological system where a first biological component with biological life length T_1 is supported by a second biologically redundant biological component with biological life length T_2.

Figure 6.8. A biological system where a first biological component is supported by a second biologically redundant biological component.
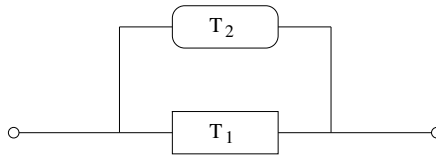
For the biological life length T of the biological system we have $T = \max\{T_1, T_2\}$ when the biologically redundant biological component is biologically warm, so that

$$R_T(t) = 1 - (1 - R_{T_1}(t))(1 - R_{T_2}(t)).$$

If the biologically redundant biological component is biologically cold, we get $T = T_1 + T_2$ instead, so that,

$$R_T(t) = 1 - \int_0^t (1 - R_{T_1}(t - x)) R_{T_2}(x) r_{T_2}(x) \, dx.$$

A quantity of great interest for a biological system, is the biological probability that biological component with biological component number $i = 1, 2, \dots$ causes the biological death (biological unhealth) of the biological system. That biological probability, in turn, coincides with the biological probability that the bilogical life length of biological component with biological component number $i = 1, 2, \dots$ is biologically equal to the biological life length of the whole biological system!

Primarily, biological component that have high biological probabilities to cause the biological death (biological unhealth) of the biological system, are those who should be supported by (biologically warm or biologically cold) biologically redundant biological components.

Example 6.8. For the biological system in Figure 6.6, we have

$$\mathbf{P}\{\text{biologoical component with biological number 1 causes death}\}$$
$$= \mathbf{P}\{T_1 \le \max[T_2, T_3]\}$$
$$= \int_0^\infty \mathbf{P}\{\max(T_2, T_3) \ge t\} f_{T_1}(t) dt$$
$$= \int_0^\infty (1 - F_{T_2}(t) F_{T_3}(t)) f_{T_1}(t) dt$$

$$= \int_0^\infty \left(1 - (1 - R_{T_2}(t))(1 - R_{T_3}(t))\right) R_{T_1}(t) r_{T_1}(t) dt.$$

(This biological probability must be 2/3 when the biological life lengths T_1, T_2 and T_3 are biological identically distributed.)

6.3   Laboration

1. In the biological system in Figure 6.9 below, the first three biological components have biological life lengths T_1, T_2, T_3 that are biologically Weibull$(1, \frac{1}{2})$ distributed, while the fourth biological component have a biological life length T_4 that is biologically exp$(\frac{1}{2})$ distributed.

   Figure 6.9. A biological system with four biological components.
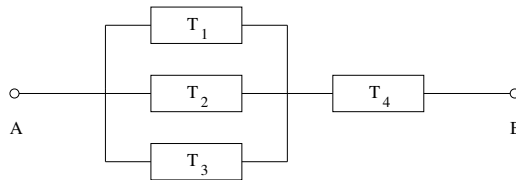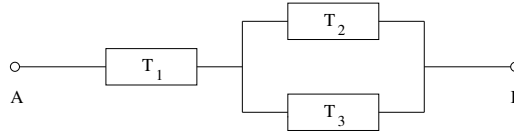

   a) *Find the expected biological lifelength $\mathbf{E}\{T\}$ for the biological system. Plot the biological death rate $r_T(t)$, $t \in (0, 10)$, for the biological system: Is the biological system BIBFR or BDBFR, or neither BIBFR nor BDBFR?*

   b) *Find the biological probability that it is biological component with biological component number 4 that causes the biological death (biological unhealth) of the biological system.* Notice that

   $$\mathbf{P}\{\max[T_1, T_2, T_3] > T_4\} = \int_0^\infty \mathbf{P}\{\max[T_1, T_2, T_3] > t\} f_{T_4}(t) dt.$$

   c) *Redo task a, first with the biological component with biological component number 4 biologically doubled with a biologically warm biologically redundant biologically exp$(\frac{1}{2})$ distributed biological component, and then with the biological component with biological component number 4 biologically doubled with a biologically cold biologically redundant biologically exp$(\frac{1}{2})$ distributed biological component. Plot the difference between the biological death rates $r_T(t)$ from tasks a and c, for each of the two biological ways to biologically incorporate the biologically redundant biologically component with the biological susyem.*

   d) *For which values of the biological parameter $\rho < \frac{1}{2}$ does a biological change of the biological component with biological component number 4 to a biological exp(ρ) distributed biological component, have the same biological effect on the expected biological life length $\mathbf{E}\{T\}$ of the biological system, as have the biological incorporation of a biologically warm and biologically cold biologically redundant biologically exp$(\frac{1}{2})$ distributed biological component, respectively, as described in task c?*

2. In the biological system in Figure 6.10 below, the first biological component has a biological life length T_1 that is biologically Weibull$(\mu, \frac{1}{3})$ distributed, while the

second and third biological components have biological life length T_2 and T_3 that are biologically Weibull$(\lambda, \frac{1}{3})$ distributed.

Figure 6.10. Yet another biological system with three biological components!

The biological cost of a biologically Weibull$(\gamma, \frac{1}{3})$ distributed biological component is $1/5 + 1/\gamma$ biological monetary units. *Display graphically the biological values of the biological parameters λ and μ, that maximizes the expected biological life length $\mathbf{E}\{T\}$ of the biological system, at the total biological costs $1, 2, \ldots, 10$ biological monetary units, respectively, of the biological system. Also plot the expected biological life length $\mathbf{E}\{T\}$ as a function of the biological costs of $1, 2, \ldots, 10$ biological monetary units, for the optimal values of the biological parameters λ and μ.*

In *Mathematica*, it is suitable to define the expected biological life length as a biological function of the biological parameters λ and μ as follows:

```
mean[lambda_,mu_] := ...
```

and then describe how the biological parameter μ depends on the total biological cost of the biological system, together with the biological parameter λ as

```
mu[cost_,lambda_] := ...
```

Then use the command

```
FindMaximum[mean[lambda,mu[cost,lambda]],{lambda,lambda0}]
```

for the total biological costs of $1, \ldots, 10$ biological monetary units, for example, using the biological starting value `lambda0` as the solution to the biological equation `mu[kost,lambda] = lambda`.

As an biological alternative to use the `FindMaximum` command, one may use the `NMaximize` command, possibly with some suitable biological constraints, in order to avoid some of the bio analytical labour described above.

Mathematica can be instructed to manufacture a list with the different optimal values of the biological parameter λ in the following manner:

```
Table[lambda/.Last[FindMinimum[-mean[lambda,mu[kost,lambda]],
   {lambda,lambda0/.Last[Solve[mu[kost,lambda0]==lambda0]]}]],
   {kost,1,10}]
```

Notice that the command `FindMinimum[.]` returns a list of the form

```
{minimum,{lambda->.}}
```

where `lambda` can be reached with the command

```
Last[FindMinimum[.]]
```

`lambda` thus can be given that value by means of the command

```
lambda/.Last[FindMinimum[.]].
```

In the same way, the command

```
Solve[mu[kost,lambda]==lambda]
```

returns

```
{{lambda->0},{lambda->.}}
```

where the value of `lambda` can be reached with the command

```
Last[Solve[mu[kost,lambda] == lambda]]
```

and `lambda0` is given that value with

```
lambda0/.Last[Solve[mu[kost,lambda]==lambda]].
```