CONSTRUCTION OF NEW MULTIVARIATE PROBABILITY DISTRIBUTIONS WITH BIOMEDICAL APPLICATIONS

Jerzy Filus
Department of Mathematics and Computer Science
Oakton Community College, Des Plaines, IL 60016, USA
email: yfilus@oakton.edu

Lidia Filus
Department of Mathematics
Northeastern Illinois University, Chicago, IL 60625, USA
email: L-Filus@neiu.edu

NORTHEASTERN ILLINOIS UNIVERSITY
MATHEMATICAL SCIENCES SEMINAR
FEBRUARY 16, 2007
The method of construction of multivariate probability distributions, presented below, is relatively new.

Initial results related to this presentation are published in:


A short communication on the initial version of results can also be found in


Continuous Multivariate Distributions.


Application of the methods for Weibullian, gamma, and lognormal distributions are presented in

Abstract: A new interesting class of multivariate probability distributions is constructed and applied in a biomedical framework. The practical objective is to determine a stochastic model that is expected to improve (in the sense of accuracy) predictions regarding patients’ (by nature random) times from a current to the next attack of illness (or to death). The measures of the predictions can be expressed as either the considered time’s expected value or, whenever possible, approximations of the whole conditional probability distribution functions of that time. The basic mathematical tool applied towards this goal, is the use of the conditional hazard rate of the considered time (given numerical values), which characterizes some essential factors that may influence the health of a patient. For example, it may be the amount and intensity of consumed tobacco and/or other potentially harmful substances. Also, one may consider amounts of specific medication administered during the process of treatment. The combined effect of these factors is discussed as well.

The formal description of the stochastic dependences between the considered time to event, and the measurable random stresses is established.
**Basic Definition:**

Suppose $X$ is a random variable that has the probability density function (pdf) $f(x)$.

By the **hazard rate** $r(x)$ of $X$ at $x$ we mean the following limit:

$$r(x) = \lim_{h \to 0^+} \frac{Pr(x \leq X < x + h \mid X \geq x)}{h}$$

$$= \frac{f(x)}{1 - F(x)} = \frac{f(x)}{Pr(X \geq x)}, \text{ for every argument } x.$$

In other words, for each small enough value of $h$ ($h > 0$) formula (A) ‘produces’ the following (linear in $h$) approximation of the conditional probability:

$$Pr(x \leq X < x + h \mid X \geq x) \cong r(x) h, \quad (as \ h \to 0).$$

The hazard (failure) is a constant if and only if the probability distribution is **exponential**.
1. PROPOSED BIOMEDICAL APPLICATIONS

1) From now on our objective is to model stochastic dependences of some random quantities, say $Y$, having a proper "medical meaning" from some random explanatory variables, here denoted by $Z_1, \ldots, Z_k$, $k = 1, 2, \ldots$. It always will be assumed that any values $z_1, \ldots, z_k$ the explanatory variables $Z_1, \ldots, Z_k$ take on, in some way influence the probability distribution of the r. variable (of interest) $Y$.

This (stochastic) influence will rely on proper changes in the values of the hazard rate, related to the probability distribution of $Y$. Speaking more generally, the values $z_1, \ldots, z_k$ determine numerical values of the parameters of the hazard rate function associated with the quantity $Y$. 
2) Now we will concentrate on modeling the following “bio-situation”.

The nonnegative random variable $Y$ will be interpreted as “time to an event”. More precisely, $Y$ models a (random) length of the time between a recovery and a next strike of illness (or to death).

This time (to be “predicted” in the sense of its probability distribution or merely its expected value) may stochastically depend on such factors like kind and amount of various stresses, time and intensity of smoking tobacco, since when the patient quit smoking or drinking.

Moreover, the foregoing phenomena while resting, so that a stress impact decay, might probably be modeled by the so called “forgetting factors” that are at our disposal as
proper mathematical tools), time and an intensity of drinking or other chemical abuse, the length of time spent in prison or other harmful condition in the past, since when a traumatic situation terminated, and many other harmful events. In addition one may consider age, blood group, level of hormones, etc.

Along with the above, some of the explanatory variables $Z_1, \ldots, Z_k$ may as well describe factors that have a positive impact on the hazard rate of the time $Y$, making that time statistically longer.

The last factors may be associated with the (proper) doses of “good” medication in the treatment processes. Needless to say, the effect of a given level of an administered drug is random.
Mathematically the problem mainly relies on a good enough description (or “definition”) of the stochastic dependences of the time to event $Y$, from its explanatory variables $Z_1, \ldots, Z_k$.

This kind of dependence seems to be very general in terms of its applicability to a wide range of real world problems.

To simplify the definition or the construction of the dependence (in terms of the mathematical model) as an illustration (or “demonstration” ) we have chosen a process of modeling the reliability of a multi-component system, by constructing the stochastic dependence as a dependence between the system’s component life-times.

This will be done next.
2. Reliability Illustration of the Problem

2.1 Given is a parallel system that consists of two components, say \( e_1, e_2 \).

By the **stochastic model of** the system’s **reliability** we will mean a joint probability distribution of the system component life-times \( X_1, X_2 \).

The goal is to describe a procedure that enables us to construct such bi-variate probability distributions as the (reliability) models.

The procedure basically consists of two stages.

**In the first stage** the two components \( e_1, e_2 \) are tested in separation of each other in special “laboratory (“off-system”) conditions”. It is assumed that, as a result of applying some common statistical methods, one obtains a good enough estimation of probability distributions \( F_1(t_1), F_2(t_2) \) of the component life-times, say \( T_1, T_2 \), respectively.
As a result of the components separation, at this stage of the procedure, the life-times $T_1$, $T_2$ are stochastically independent.

In the second stage the components are assumed to be installed into the real system, where interactions between the components take place. In other words, some (additional) physical phenomena, associated with a component functioning, contribute to the failure mechanism of the other.

[ As examples of such “physical phenomena” one may encounter: an increase in temperature, rising level of mechanical vibrations, or other unavoidable side-effects that follow the activities associated with operating of one of the components, say $e_1$. ]

Therefore, compared to the “original” laboratory conditions, additional “stresses” are put on at least one of the system’s components when in-system.

(At this point we adopt the simplifying assumption that ‘physically’ component $e_1$ influences component $e_2$, but $e_1$ is not influenced (“back”) by $e_2$.)
This additional stress (the physical impact) is thought of as a sequence of “many” “small” micro-shocks by which $e_1$ is constantly “bombarding” $e_2$. The sequence of these micro-shocks produced by $e_1$ creates a corresponding sequence of micro-damages in component $e_2$, affecting its reliability (or life-time). The above physical phenomena, often too vague and too complicated to be handled directly (in terms of pure physical theories) has the following, much simpler stochastic equivalence.

In a “translation” from physical to stochastic description, the micro-damages created in the physical structure of $e_2$ are assumed to have their corresponding “representations” as changes, of a proper magnitude, in the hazard rate of component $e_2$. Consequently, its (new) life time, say $X_2$, becomes statistically longer or shorter than the (original) life time $T_2$, when in the laboratory conditions.

(In short, we have a transformation $T_2 \rightarrow X_2$, while $T_1 = X_2$, in distribution.)
In order to grasp the somewhat vague, poorly recognized, physical interactions between the system components in the stochastic reliability model, the following will be assumed:

1) The “magnitude” or an “amount of influence” (from a reliability point of view) of \( e_1 \) on \( e_2 \), among other things, depends explicitly on the length of time \( X_1 \) of the components’ common interaction, assuming that the random event \( X_1 < X_2 \) occurs with probability close to 1.

The stochastic dependence (of \( X_2 \) from \( X_1 \)) will be described by a suitable continuous function of the time \( x_1 \) the impact of \( e_1 \) on \( e_2 \) lasted (i.e., whenever the event \( X_1 = x_1 \) happens), in the following way:
2) Suppose component $e_1$ fails before a failure of $e_2$ (i.e., $X_1 < X_2$). Then the original hazard rate, say $\lambda_2(x_2; \theta_2)$, of component $e_2$ that corresponds to the original pdf $f_2(x_2, \theta_2)$ of the laboratory conditions life time $T_2$, is subjected to a change in value. This change of value of the hazard rate $\lambda_2(x_2; \theta_2)$ of $e_2$ is assumed to be reflected by a change in its (scalar or a vector) parameter $\theta_2$.

The magnitude of this change obviously depends on the (random) time $X_1$ the influence of $e_1$ on $e_2$ lasted. If the random event $X_1 = x_1$ happens, the new value $\theta_2^*$ of the parameter $\theta_2$ of the hazard rate $\lambda_2(x_2, \theta_2)$ clearly depends on that time $x_1$.

Therefore the new parameter $\theta_2^*$ (the hazard rate of the life time $X_2$ of $e_2$, when in system) becomes a (continuous) function of the time $x_1$. This fact justifies the functional notation:

$$\theta_2^* = \theta_2(x_1).$$

As a consequence, the hazard rate $\lambda_2(x_2, \theta_2)$, corresponding to the life time $T_2$, turns into a slightly different hazard rate $\lambda_2(x_2, \theta_2(x_1))$, which corresponds to the life time $X_2$, given that $X_1 = x_1$ happened.
Thus, we have obtained the **conditional hazard rate** $\lambda_2 (x_2 \mid x_1)$ of life time $X_2$, given $X_1 = x_1$, defined by the following formula:

$$\lambda_2 (x_2 \mid x_1) = \lambda_2 (x_2, \theta_2 (x_1)),\$$

where $\theta_2 (x_1)$ is a suitably chosen continuous function of the time $x_1$.

At this point realize that the **probability density** $f_2 (x_2, \theta_2)$ of $T_2$ is subjected to “the same” transformation of the parameter $\theta_2$ into $\theta_2 (x_1)$ as the hazard rate $\lambda_2 (x_2, \theta_2)$ is.

Thus, **in parallel** to the definition of the conditional hazard rate, given above, the formula:

$$g_2 (x_2 \mid x_1) = f_2 (x_2, \theta_2 (x_1))$$

defines the **conditional pdf** of the random variable $X_2$, given $X_1 = x_1$. 

14
[ In this way the stochastic dependence of a random variable \( X_2 \) on another r. variable \( X_1 \) was defined. ]

As consequence of the assumption, that component \( e_2 \) has no physical influence on \( e_1 \) we also have that \( X_1 = T_1 \) in distribution, so that the marginal pdf \( g_1 (x_1) \) of the life time \( X_1 \) is known to be equal to \( f_1 (x_1) \), ( upon the convention \( x_1 = t_1 \) ).

One then obtains the joint pdf \( g (x_1, x_2) \) of the random vector \( (X_1, X_2) \) of the components life times, when in system, as the usual arithmetic product:

\[
g (x_1, x_2) = g_2 (x_2 | x_1) g_1 (x_1).
\]

The above construction we illustrate by the following example:
Example 1

Suppose that the component life times in laboratory conditions are the following independent and exponentially distributed random variables $T_1, T_2$.

Let $f_k(t_k) = (1/\theta_k) \exp(-t_k/\theta_k)$

be the pdf of $T_k$ ($k = 1, 2$).

Letting $\theta_2 = \theta_2(x_1)$ (but for the rest arbitrary) in the exponential pdf $f_2(x_2; \theta_2)$, given above one obtains a wide class of bivariate (pseudo)exponential pdfs:

$$g(x_1, x_2) = g_1(x_1) \, g_2(x_2 | x_1)$$

$$= (\theta_1)^{-1} \exp[-x_1/(\theta_1)] \, (\theta_2(x_1))^{-1} \exp[-x_2 / (\theta_2(x_1))]$$

where, in particular, one may specify: $\theta_2(x_1) = \theta_2(1 + Ax_1^r)$,

(with $A, r$ positive reals).
Another analytically interesting model is obtained by setting:

$$\theta_2(x_1) = \theta_2 \exp[ A x_1^r ] ,$$

with ‘A’ being an arbitrary real ‘parameter or parameter function’.

Note that both factors $g_1(x_1)$, $g_2(x_2 | x_1)$ of $g(x_1, x_2)$, given above, are exponentials, which justifies the pdf’s name “exponential”. However, the marginal pdf $g_2(x_2)$ of $X_2$ is, in general, not an exponential.

The pattern for the above described stochastic dependence through a transformation of a pdf’s parameter into the pdf’s parameter function seems to be proper for applications in a much wider area of real world phenomena than reliability problems only, since the random variable $X_1$ present in such models may be considered as any explanatory variable for a corresponding r. variable $X_2$ of interest.

As an example consider (“any”) two random variables $X_1, X_2$ having arbitrary interpretations. Also, the interpretation of one may be quite different from the other:
Example 2
Let the random variables $T_1$, $T_2$ be independent, having the normal pdfs
\[ f_1(x_1) = N(\mu_1, \sigma_1) \text{ and } f_2(x_2) = N(\mu_2, \sigma_2) \text{ respectively.} \]

Then the corresponding bivariate “pseudonormal” pdf $g(x_1, x_2)$ of the related r. vector $(X_1, X_2)$ is given by the product formula:
\[ g(x_1, x_2) = g_1(x_1) \cdot g_2(x_2 | x_1), \]
with the invariant pdfs $g_1(x_1) = f_1(x_1) = N(\mu_1, \sigma_1)$, and with conditional pdf, given as follows:
\[ g_2(x_2 | x_1) = N(\mu_2(x_1), \sigma_2(x_1)) , \]
where the replaced parameters of $f_2(x_2)$, now are the following (arbitrary) continuous “parameter functions”
\[ \mu_2 = \mu_2(x_1), \quad \sigma_2 = \sigma_2(x_1) \text{ of the r. event } X_1 = x_1. \]

More explicitly, one obtains the following class of bivariate pseudonormal pdfs:
\[ g(x_1, x_2) = \]
\[ = [\sigma_1 \sqrt{(2\pi)}]^{-1} \exp[- (x_1 - \mu_1)^2 / 2\sigma_1^2] \ [\sigma_2(x_1) \sqrt{(2\pi)}]^{-1} \]
\[ \exp[- (x_2 - \mu_2(x_1))^2 / 2[\sigma_2(x_1)]^2], \]

where

\[ \mu_2(x_1) = E[X_2 | x_1] \]

is the (in general \textit{nonlinear}) \textit{regression function}, and

\[ [\sigma_2(x_1)]^2 = \text{Var} [X_2 | x_1] \]

is the \textit{conditional variance}, which may also be chosen to be an \textit{arbitrary} function continuous of \( x_1 \).

In particular, one may consider \textit{the (pseudonormal) nonlinear regression function} of the form:

\[ E[X_2 | x_1] = \mu_2(x_1) = \mu_2 + a (x_1 - \mu_1) + A (x_1 - \mu_1)^n, \]

with \( a \) and \( A, \) \textit{and} \( n \) being arbitrary real numbers.
For the coefficient ‘A’ (‘small’, in a comparison to ‘a’), the term $A (x_1 - \mu_1)^n$ may be considered as a nonlinear “correction” to the regular Gaussian (linear) regression function. The purpose of that correction is mainly to enhance the accuracy in various modeling situations. Specially interesting seems to be the “quadratic case” $n = 2$ which by nature is non-symmetric, contrary to the regular bi-variate normal. At this point realize the potential usefulness of the quadratic correction in situations where empirical data show a significant asymmetric tendency, while “the best” (normal) model available is a symmetric one. \[\square\]
2.2 Multivariate Extension of Bivariate Models

The bi-variate stochastic pdfs constructed so far can be extended to any arbitrary set of \( r \) variables
\[ \{X_1, X_2, \ldots, X_{n-1}, X_n \}, \quad n = 3, 4, \ldots. \]

As an illustration consider again, similar to the previous case \((n = 2)\), a parallel system whose components are \( e_1, e_2, \ldots, e_n \).

Let, as before, \( f_1(t_1; \theta_1), f_2(t_2; \theta_2), \ldots, f_n(t_n; \theta_n) \) be the pdfs of the independent component life times \( T_1, T_2, \ldots, T_n \) respectively, tested in laboratory conditions.

The stochastically dependent life times of the components, when working in the system, will be denoted by \( X_1, X_2, \ldots, X_n \).

Similarly to the bi-variate case we impose the following two conditions:

(1) for each \( j = 2, 3, \ldots, n \), components \( e_1, e_2, \ldots, e_{j-1} \) may have physical impact on component \( e_j \), while components \( e_{j+1}, \ldots, e_n \) do not. Consequently, no component has ever a physical influence on \( e_1 \),

(2) the probability of the random event \( X_1 < X_2 < \ldots < X_n \) is approximately equal to 1.

[ This pretty strong condition is dropped in most of the other applications of the constructed \( n \)-variate pdfs. ]

21
The **stochastic dependencies** between the life times \( X_1, X_2, \ldots, X_n \) of the components basically are **of the same nature** as in the bi-variate case. For each \( j = 2, 3, \ldots, n \), the way the components \( e_1, e_2, \ldots, e_{j-1} \) (when working in the system) physically and stochastically affect component \( e_j \) is quite **the same** as the way component \( e_1 \) influences component \( e_2 \) in the bi-variate model. Stochastically, their **total influence on the hazard rate** \( \lambda_j(t_j; \theta_j) \) of component \( e_j \)‘s life time \( T_j \) (in laboratory conditions), is reflected by a **change in** the value of the (scalar or vector) **parameter** \( \theta_j \). As in the bi-variate case, the magnitude of that change is assumed to depend on the (random) times \( X_1, \ldots, X_{j-1} \) the stresses put by components \( e_1, e_2, \ldots, e_{j-1} \) on \( e_j \) lasted. Consequently, the **parameter** \( \theta_j \) of the hazard rate \( \lambda_j(t_j; \theta_j) \) of the (original) life time \( T_j \) is “**replaced**” by parameter \( \theta_j^* = \theta_j(x_1, \ldots, x_{j-1}) \), of the (‘in system’) life time \( X_j \), which is **continuously** dependent on the times \( x_1, \ldots, x_{j-1} \), as the random events \( (X_1 = x_1, \ldots, X_{j-1} = x_{j-1}) \) happen.

The continuous functions \( \theta_j^* = \theta_j(x_1, \ldots, x_{j-1}) \) are called “**parameter functions**”.

Realize that these parameter functions (whenever known) determine the following **conditional hazard rates** for each component’s \( e_j \), “in-system” life-time \( X_j \):
\[ \lambda_j(x_j | x_1, \ldots, x_{j-1}) = \lambda_j(x_j ; \theta_j(x_1, \ldots, x_{j-1})) . \]  \hfill (1)

In parallel, as a result of the same transformation of the parameters \( \theta_j \) into the parameter functions \( \theta_j(x_1, \ldots, x_{j-1}) \), the pdfs \( f_j(t_j ; \theta_j) \) of \( T_j \)‘s are transformed into the conditional pdfs \( g_j(x_j | x_1, \ldots, x_{j-1}) \) of the ‘in-system’ life times \( X_j \), given the random events \( (X_1 = x_1, \ldots, X_{j-1} = x_{j-1}) \) took place.

The conditional pdfs are then defined by the following sequence of the identities:

\[ g_j(x_j | x_1, \ldots, x_{j-1}) = f_j(x_j ; \theta_j(x_1, \ldots, x_{j-1})) , \]  \hfill (2)

as \( j = 2, 3, \ldots, n \).

The above formula provides a full description of the stochastic dependencies among the component ‘in system’ life times \( X_1, \ldots, X_n \).

Also realize that the mathematical descriptions (1) and (2) of the phenomena are equivalent to each other.
The **general pattern** for the joint pdfs of the random vectors 
\( (X_1, \ldots, X_n) \) construction, for \( n = 1, 2, 3, \ldots \), may be thought of as a result of the following **recurrence procedure**. This procedure contains the following steps:

1. For \( n = 1 \), the “original” marginal pdf \( g_1(x_1) \) of \( X_1 \) is given by the assumption that \( g_1(x_1) = f_1(x_1) \), where \( f_1(x_1) \) is an “arbitrary” ‘off-system’ pdf of \( T_1 \), given in advance.

2. For \( n = 2 \), the bi-variate pdf \( g(x_1, x_2) \) of the r. vector \((X_1, X_2)\) one obtains by the method described in section 1. (Recall that it was obtained in the factored form

\[
g(x_1, x_2) = g_1(x_1) g(\ x_2 \mid x_1 \ ) .
\]

3. For \( n \geq 3 \), we proceed as follows.
   If, for some \( j = 3, 4, \ldots, n \), the **joint pdf** \( g^{j-1}(x_1, \ldots, x_{j-1}) \) of the random vector \((X_1, \ldots, X_{j-1})\) is obtained as the result of the \((j-1)^{th} \) **iteration**, then the \( j^{th} \) **iteration** yields the \( j^{th} \) joint pdf \( g(x_1, \ldots, x_j) \) of the r. vector \((X_1, \ldots, X_j)\) as the usual arithmetic product:

\[
g^j(x_1, \ldots, x_j) = g^{j-1}(x_1, \ldots, x_{j-1}) \ g_j(\ x_j \mid x_1, \ldots, x_{j-1} \ ) ,
\]

24
where the conditional pdf \( g_{j}(x_{j} \mid x_{1}, \ldots, x_{j-1}) \) is given by formula (2) on the previous page, for a given continuous parameter function \( \theta_{j}(x_{1}, \ldots, x_{j-1}) \).

4) Obviously the procedure stops when \( j = n \).
(Otherwise the stochastic process is well defined.)

As a final result one obtains the **stochastic reliability model** as the joint pdf \( g(x_{1}, \ldots, x_{n}) \) of the component ‘in-system’ life times \( X_{1}, \ldots, X_{n} \).

Notice that from the above recurrence procedure description it follows that for each \( n = 2, 3, \ldots \) the obtained joint pdf \( g(x_{1}, \ldots, x_{n}) \) can always be represented as the **product** of exactly \( n \) **factors**:

\[
g(x_{1}, \ldots, x_{n}) = g_{1}(x_{1}) g_{2}(x_{2} \mid x_{1}) g_{3}(x_{3} \mid x_{1}, x_{2}) \ldots g_{n}(x_{n} \mid x_{1}, \ldots, x_{n-1}) \quad (3)
\]
Example 3

To illustrate the general n-variate case assume that the \( r \) variables \( T_1, \ldots, T_n \) are independent, each having, for \( i = 1, \ldots, n \), the following **Weibull** pdf:

\[
f_i(t_i) = \left( \frac{\gamma_i}{\beta_i} \right) (t_i - \alpha_i)^{\gamma(i)-1} \exp\left[ - (t_i - \alpha_i)^{\gamma(i)} / \beta_i \right], \quad t_i > \alpha_i, \\
= 0 \text{, elsewhere.}
\]

We write \( f_i(t_i) = W_i(\alpha_i; \beta_i, \gamma(i)) \) and \( \gamma_i = \gamma(i) \).

Applying the pattern of the parameter replacement method (in a simple procedure shown below) one obtains the class of **Weibullian** joint pdfs \( g(x_1, \ldots, x_n) \) of the \( r \) vectors \( (X_1, \ldots, X_n) \) in the product form (3). According to the procedure, we first set

\[
g_1(x_1) = f_1(x_1) = W_1(\alpha_1, \beta_1, \gamma(1)).
\]

Next, for each \( j = 2, \ldots, n \), the conditional pdf \( g_j(x_j | x_1, \ldots, x_{j-1}) \) of \( X_j \) becomes the \( j^{th} \) factor in product (3). It is noteworthy that every such factor remains **Weibullian**, each one with respect to \( x_j \) alone.

The conditional pdf is given by the following specification of the (three parameter **Weibullian**) parameter function:
\[ g_j(x_j | x_1, \ldots, x_{j-1}) = W_j(\alpha_j(x_1, \ldots, x_{j-1}) \ ; \ \beta_j(x_1, \ldots, x_{j-1}) \ ; \ \gamma_j(x_1, \ldots, x_{j-1}) ) . \]  

(4)

“Theoretically” the continuous parameter functions \( \alpha_j(x_1, \ldots, x_{j-1}) \), \( \beta_j(x_1, \ldots, x_{j-1}) \), and \( \gamma_j(x_1, \ldots, x_{j-1}) \) can be declared in a, basically, arbitrary way. They represent the Weibullian parameters of the shift, the scale, and the shape respectively. Each of them continuously depends on occurring random events \( (X_1, \ldots, X_{j-1}) = (x_1, \ldots, x_{j-1}) \). Once used as stochastic models (of system reliability, for example) the ‘arbitrariness’ of the parameter functions must obviously be replaced by a rather careful choice of “the best”. For such a choice use of statistical methods of estimation and discrimination is necessary, when modeling.
Among all the models considered in this paper, the most essential are the conditional pdfs $g_j(x_j \mid x_1, \ldots, x_{j-1})$ that occur in a variety of different contexts, regardless of the pdf’s class they belong to.

It seems that the range of applications of this may in fact be very wide. This range includes, among others, problems of actuary and many of the bio-medical problems.

This fact becomes more apparent if we abstract the meaning of the r. variables $X_1, \ldots, X_{j-1}, X_j$ used throughout this presentation ($j = 2, 3, \ldots$).

Accordingly, the common (in the constructed reliability models) expression $g_j(x_j \mid x_1, \ldots, x_{j-1})$ from now on will be replaced by a more general expression, say $g(y \mid z_1, \ldots, z_k)$. Actually, the two expressions are mathematically identical, but when modeling non-mathematical ‘realities’, the practical meaning of the random variables, now denoted by $Y, Z_1, \ldots, Z_k$, (with $k = j-1$) is not restricted anymore.

In particular, $Y$ may also be considered as a random vector.

However, the random variables $Z_1, \ldots, Z_k$ will be assumed to serve as explanatory variables for the $Y$. 
Besides, no restrictions need be imposed on the probability distributions classes the r. variables $Y, Z_1, \ldots, Z_k$ belong to.

The only requirement that is essential is pre-knowledge of the joint pr. distribution of the random vector $(Z_1, \ldots, Z_k)$.

However, as a first approach the r. variables $Z_1, \ldots, Z_k$ (that have given univariate pdfs) may be considered independent.
5. We now return to our original **bio-medical problem** of predicting the random time $Y$ that runs from a moment of recovery to the next attack of the illness or to death.

The prediction is to be based on the possible knowledge of $Y$’s **conditional pdfs**, given numerical values of the explanatory factors $Z_1, \ldots, Z_k$ that should provide additional information (comparing to those used to a date) on inner and outer (“environmental”) conditions “surrounding” the patient.

As an effect of making use of **more information than is currently done**, one expects to achieve a significantly **better prediction** of the random time $Y$’s value.

An expected enhancement of the prediction accuracy should be the result of having at hand a more precise probability distribution of the quantity (the time) of interest.
TWO EXAMPLES OF WEIBULL STOCHASTIC MODELS

In the model that follows the initial pdf (in absence of the stresses, i.e., in “laboratory conditions”, so that all explanatory values are equal zero), say $f_1(y)$, of the ‘time to event’ $Y$ is given as follows:

$$f_1(y) = \lambda y^{\gamma-1} \exp\{ -[\lambda y^{\gamma}] \},$$

and $r(y) = \lambda y^{\gamma-1}$, is the corresponding hazard rate.

When the stresses expressed in terms of the explanatory variables occur so that when the r. event $(Z_1 = z_1, Z_2 = z_2, \ldots, Z_k = z_k)$ happens, we assume that these will affect the (original) hazard rate through the scale parameter $\lambda$. As a result, this parameter is transformed into a ‘parameter function’ as follows:

$$\lambda \rightarrow \lambda \left(1 + a_1 z_1^\beta + a_2 z_2^\beta + \ldots + a_k z_k^\beta \right),$$

while in this particular example we keep the shape parameter ‘$\gamma$’ invariant. [This last requirement is, however, not necessary]
As a result we obtain the transformation for the Weibull hazard rates:

\[ r(y) \rightarrow R(y \mid z_1, \ldots, z_k), \] or

\[ \lambda y^{\gamma^{-1}} \rightarrow \lambda \left(1 + a_1 z_1^\beta + a_2 z_2^\beta + \ldots + a_k z_k^\beta\right) y^{\gamma^{-1}}. \]

The same operation can also be expressed in terms of the pdfs transformation: \( f_1(y) \rightarrow g_1(y \mid z_1, \ldots, z_k) \), i.e.,

\[ f_1(y) = \lambda y^{\gamma^{-1}} \exp\left\{-\left[\lambda y^{\gamma}\right]\right\} \]

\[ \rightarrow g_1(y \mid z_1, \ldots, z_k), \]

where the latter conditional pdf is the stochastic model for the time \( Y \) to event given by the following formula:

\[ g_1(y \mid z_1, \ldots, z_k) \]

\[ = \left[\lambda \left(1 + a_1 z_1^\beta + a_2 z_2^\beta + \ldots + a_k z_k^\beta\right)\right] y^{\gamma^{-1}} \]

\[ \exp\left\{-\left[\lambda \left(1 + a_1 z_1^\beta + a_2 z_2^\beta + \ldots + a_k z_k^\beta\right)\right] y^{\gamma}\right\}, \] (6)
where all the coefficients $a_i$ ($j = 1, \ldots, k$) in (6) are positive, and each of them may depend on the choice of treatment, given the type of illness.

Another Weibullian model, alternative to (6), can also be applied. Assume the following transformation of the scale parameters:

$$
\lambda \rightarrow \lambda \exp[b_1 z_1 \beta + b_2 z_2 \beta + \ldots + b_k z_k \beta].
$$

This produces the following Weibull (in argument ‘$y$’) conditional pdf:

$$
g_2(y | z_1, \ldots, z_k) = \lambda \exp[b_1 z_1 \beta + b_2 z_2 \beta + \ldots + b_k z_k \beta] \ y^{\gamma-1} \ \exp[-\lambda \exp[b_1 z_1 \beta + b_2 z_2 \beta + \ldots + b_k z_k \beta] \ y^\gamma], \quad (7)
$$

where the coefficients $b_i$ ($i = 1, \ldots, k$) are “arbitrary” (possibly also negative), and may depend on the particular choice of treatment.
**Remark:** Realize that in both the conditional pdfs (6) and (7) only the dependence of the time $Y$ from each particular amount of (one) stress, say $z_j$ ($j = 1, 2, \ldots, k$) is taken into consideration. But from everyday practical experience it is well known that often one encounters effects that are results of some combination of more than one stress. As it often happens neither one of the stresses taken separately could cause a given additional stress. For example, a chemical interaction between two or more medications may yield a stress quite different than each of the medications alone would produce.

The simplest way to describe these combined effects in a stochastic model such as (6) or (7) seems to be to include in the scale parameter functions terms that are monotonic functions of expressions like the products $(z_i z_j)$, $i, j = 1, \ldots, k$. These terms then imply some stochastic correspondence to the joint effects of $i^{th}$ and $j^{th}$ stress.
For example, the scale parameter function in formula (6) could be extended to the following expression:

\[ \lambda \exp\left[ b_1 z_1^\beta + b_2 z_2^\beta + \ldots + b_k z_k^\beta \right] \]

\[ \Rightarrow \lambda \exp\left[ b_1 z_1^\beta + b_2 z_2^\beta + \ldots + b_k z_k^\beta + \sum c_{ij} (z_i z_j)^\beta \right] , \]

with, possibly small, coefficients \( c_{ij} \) (most of them zeros).

Similarly, the joint effect of three factors could be described by functions of the products \( (z_i z_j z_m) \) etc … .