Scheduling Independent Tasks to Minimize the Makespan on Identical Machines

John Bruno[1]
AT&T Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

Edward G. Coffman, Jr.
AT&T Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

Peter Downey
Department of Computer Science
University of Arizona
Tucson, AZ 85721

TR 94-13

Abstract

In this paper we consider scheduling n tasks with task times that are i.i.d. random variables with a common distribution function F on m parallel machines. Scheduling is done by an *a priori* assignment of tasks to processors. We show that if the distribution function F is a Pólya frequency function of order 2 then the assignment which attempts to place an equal number of tasks on each processor achieves the stochastically smallest makespan among all assignments. The condition embraces many important distributions, such as the gamma and truncated normal distributions.

November 1, 1993

Department of Computer Science
The University of Arizona
Tucson, AZ 85721

1. Currently on sabbatical from the Department of Computer Science, University of California, Santa Barbara. The work of this author was partially supported by a UC MICRO Grant and by the Xerox Corporation.

Scheduling Independent Tasks to Minimize the Makespan on Identical Machines

John Bruno[1]
AT&T Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

Edward G. Coffman, Jr.
AT&T Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

Peter Downey
Department of Computer Science
University of Arizona
Tucson, AZ 85721

November 1, 1993

Abstract

In this paper we consider scheduling n tasks with task times that are i.i.d. random variables with a common distribution function F on m parallel machines. Scheduling is done by an *a priori* assignment of tasks to processors. We show that if the distribution function F is a Pólya frequency function of order 2 then the assignment which attempts to place an equal number of tasks on each processor achieves the stochastically smallest makespan among all assignments. The condition embraces many important distributions, such as the gamma and truncated normal distributions.

1.0 Introduction

In this paper we consider scheduling $n \geq 1$ tasks on $m \geq 2$ identical machines. The task processing times are non-negative, independent, identically distributed random variables. Scheduling is done by an *a priori* assignment of the tasks to the machines. Our objective is to find an assignment which minimizes the makespan, the latest finishing time among all the tasks. We seek an assignment under which the makespan is stochastically majorized by that of any other assignment. One might expect that the assignment which places as nearly as possible the same number of tasks on each machine would give the stochastically smallest makespan. In the following example we show that this is not necessarily the case.

Example: Let $n = 4$ and $m = 2$. The distribution function for the task processing times is: $F(t) = 0$ for $t < 2$; $F(t) = p$ $(0 < p < 1)$ for $2 \leq t < 5$; and $F(t) = 1$ for $t \geq 5$. Assignment A places two tasks on each machine and assignment B places three tasks on one

1. Currently on sabbatical from the Department of Computer Science, University of California, Santa Barbara. The work of this author was partially supported by a UC MICRO Grant and by the Xerox Corporation.

machine and one task on the other. Let M_A and M_B denote the makespan of assignments A and B, respectively. It is easy to see that $Pr\{M_A \le 6\} = p^4$ and $Pr\{M_B \le 6\} = p^3$. The reason is that under assignment A all four tasks must have processing time equal to 2 if all of the tasks are to finish by time 6. Under assignment B the processing time of the lone task on one of the machines can be either 2 or 5 and the remaining three tasks on the other machine all have to have processing time 2 in order for the makespan to be no larger than 6. The makespan for assignment A is not stochastically smaller than the makespan under B. □

In what follows we show that if the distribution function of the task processing times is a Pólya frequency function of order 2 then the "flattest" assignment of tasks to processors stochastically minimizes the makespan. It is known that this assignment minimizes the makespan in the sense of convex ordering for any distribution function [Chang 1992].

2.0 Main Results

Before stating our results we need a few definitions. An *assignment* π is an m-vector of nonnegative integers such that the sum of its components equals n. Let $\pi = (\pi_1, \ldots, \pi_m)$ and $\pi_{[1]} \ge \ldots \ge \pi_{[m]}$ denote the components of π in decreasing order. Let π and π' be two assignments. We say π is *majorized* by π' [MO 1979] if

$$\sum_{i=1}^{k} \pi_{[i]} \le \sum_{i=1}^{k} \pi'_{[i]} \quad \text{for } k = 1, \ldots, m-1.$$

Since π and π' are assignments their component sums are both equal to n.

Let X and Y be totally ordered sets. A function $K(x, y)$ on $X \times Y$ is *totally positive of order 2* (TP_2) if

$$\begin{vmatrix} K(x_1, y_1) & K(x_1, y_2) \\ K(x_2, y_1) & K(x_2, y_2) \end{vmatrix} \ge 0$$

for all $x_1 \le x_2$ in X and $y_1 \le y_2$ in Y. A non-negative function $h(x)$ on $R = (-\infty, \infty)$ is a *Pólya frequency function of order 2* (PF_2) if $K(x, y) = h(x - y)$ is TP_2 on $R \times R$ [Karlin 1968].

Let $F(t)$ denote the distribution function of the task times. Throughout we assume that the distribution function of the task times is such that there exists a non-negative, continuous function $f(t)$ such that for all t we have $F(t) = \int_{-\infty}^{t} f(x)dx$.

If π is an assignment, then M_π denotes the makespan random variable under the assignment π. Let X and Y be random variables with distribution functions $F_X(t)$ and $F_Y(t)$, respectively. We say that X is *stochastically less than* Y, denoted $X \le Y$, if $F_X(t) \ge F_Y(t)$ for all t. We use the symbol $\le$ to denote the stochastic ordering relation between random variables and the usual numerical ordering relation. In what follows it will be clear from the context which ordering is intended. An alternative and useful characterization of stochastic dominance is the following: X is *stochastically less than* Y if and only if for all increasing functions g we have $E[g(X)] \le E[g(Y)]$, whenever the expectations exist [Stoyan 1983].

The next section proves the following result.

Theorem A. *If $F(t)$ is PF_2 then $M_\pi \le M_{\pi'}$ whenever π is majorized by π'.*

It follows from Theorem A that an assignment π which distributes the tasks as evenly as possible among the processors (the maximum difference between the number of tasks on any two processors is at most one) is majorized by all other assignments. Accordingly, M_π is stochastically less than the makespan of any other assignment.

In the following Theorem we assume there are $k(k+1)$ tasks and compare the makespan of the best assignment on $k+1$ machines with that of the best assignment on k machines.

Theorem B. *The makespan of the assignment which places k tasks on each of $k+1$ machines is stochastically less than the makespan of the assignment which places $k+1$ tasks on each of k machines.*

Proof: The following appears in [BP 1975]. Every distribution function F such that $F(t) = 0$ for $t < 0$ has the property that $[F^{(k)}(t)]^{1/k}$ is decreasing in $k = 1, 2, \ldots$, where $F^{(k)}(t)$ denotes the k-fold convolution of $F(t)$ with itself. Using this result it follows that $[F^{(k)}(t)]^{k+1} \ge [F^{(k+1)}(t)]^k$ for all positive integers k. □

3.0 Proof of Theorem A

Suppose π is majorized by π'. Then there exits a sequence of assignments $\pi^1, \ldots, \pi^k$ such that $\pi' = \pi^1$, $\pi = \pi^k$, π^i majorizes π^{i+1} for $i = 1, \ldots, k-1$, and consecutive assignments differ in exactly two components. In particular, we can restrict ourselves to the following operation in going from one assignment to the next: subtract one from one component and add one to some other component where the first component is greater than the second component prior to the operation.

Let A_i and B_i denote the makespans under assignments π^i and π^{i+1}, respectively, corresponding to the two machines whose assignments are changed in going from assignment π^i to π^{i+1}. Let R be the makespan of the tasks on the remaining $m-2$ machines. Then

$$M_{\pi^{i+1}} = max(B_i, R)$$

and

$$M_{\pi^i} = max(A_i, R) .$$

Lemma 1. If $B_i \le A_i$ then $M_{\pi^{i+1}} \le M_{\pi^i}$.

Proof: This follows since $F_{B_i} F_R \ge F_{A_i} F_R$. □

Lemma 2. Let Y and Z be independent random variables and x be a nonnegative real. Then $max(Y+x, Z) \le max(Y, Z+x)$ if and only if $F_Y(t-x)F_Z(t) \ge F_Z(t-x)F_Y(t)$.

Proof: Immediate. □

Lemma 3. Suppose X is distributed according to $F(t)$ and $max(Y+x, Z) \le max(Y, Z+x)$ for all $x \ge 0$. Then $max(Y+X, Z) \le max(Y, Z+X)$.

Proof: Show that $E[g(max(Y+X, Z))] \le E[g(max(Y, Z+X))]$ for all increasing functions g by conditioning on X. □

Let $F^{(k)}(t)$ denote the k-fold convolution of $F(t)$ with itself for $k \ge 0$, where $F^{(0)}(t) \equiv 1$ and $F^{(k)}(t) = \int F^{(k-1)}(t-x)f(x)dx$ for $k \ge 1$. The limits of integration are taken from $-\infty$ to $+\infty$ unless stated otherwise.

Theorem 1. If $F(t)$ is PF_2 then $F^{(k)}(t)$ is TP_2 in k and t.

Proof: We give a proof of this result in the Appendices. This theorem appears as problem 5 on page 98 in [BP 1975]. □

Lemma 4. If Y and Z are distributed according to $F^{(n_1)}(t)$ and $F^{(n_2)}(t)$, respectively, where $n_1 \le n_2$ and X is distributed according to $F(t)$ then $max(Y + X, Z) \le max(Y, Z + X)$.

Proof: The lemma follows from Theorem 1 and Lemmas 2 and 3. □

Theorem A follows from Lemmas 1 and 4.

4.0 Conclusions

In this paper we show that even with independent and identically distributed task times the *a priori* assignment which places as nearly as possible an equal number of tasks on each machine does not necessarily yield the stochastically minimum makespan over all assignments. We have provided a sufficient condition on the distribution function of the task time random variable which guarantees that "flatter" assignments are stochastically better assignments and, in particular, that the "flattest" assignment is stochastically minimum over all assignments.

We also show that for any distribution function and $n = k(k+1)$ that assigning k tasks to each of $k+1$ machines yields a makespan which is stochastically smaller than the makespan achieved by assigning $k+1$ tasks to each of k machines.

It would be interesting to characterize the class of distribution functions for which $M_\pi \le M_{\pi'}$ if and only if π is majorized by π'. It is known that for any distribution function M_π is less than $M_{\pi'}$ in the sense of convex ordering of random variables whenever π is majorized by π' [Chang 1992][MN 1993].

5.0 References

[BP 1965] Barlow, Richard E., and Frank Proschan, *Mathematical Theory of Reliability*, John Wiley, New York, 1965.

[BP 1975] Barlow, Richard E., and Frank Proschan, *Statistical Theory of Reliability and Life Testing: Probability Models*, Holt, Rinehart and Winston, New York, 1975.

[Chang 1992] Chang, C. S., A New Ordering for Stochastic Majorization: Theory and Applications, *Advances in Applied Probability*, 24 (1992), pp. 604-634.

[EMP 1973] Esary, J.D., A.W. Marshall, and F. Proschan, Shock Models and Wear Processes, *The Annals of Probability*, Vol. 1, No. 4, 1973, pp. 627-649.

[Karlin 1968] Karlin, Samuel, *Total Positivity*, Volume 1, Stanford University Press, 1968.

[MN 1993] Makowski, A. M. and R. Nelson, Characterizing Majorization on $\mathbb{N}^K$, with Applications to Multiprocessor Scheduling, *Technical Report ISR 93-227*, Institute for Systems Research, University of Maryland, College park, MD, 1993.

[MO 1979] Marshall, Albert W. and Ingram Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, 1979.

[Stoyan 1983] Stoyan, D., *Comparison Methods for Queues and Other Stochastic Models*, New York, John Wiley, 1983.

Appendix A: PF_2 Distributions are Closed Under Convolution

It is shown in [BP 1965] that if $\bar{F}_1$ and $\bar{F}_2$ are PF_2 then $\bar{H}$, where H is the convolution of F_1 and F_2, is PF_2. In this appendix we give a proof of the assertion that if F_1 and F_2 are PF_2 then H is also PF_2. Our proof is patterned after the proof in [BP 1965] of the former assertion. While the closure property of PF_2 distributions is often cited in the literature, we have not been able to find a direct proof of this result. See [BP 1965] Theorem 5.3.

The following definitions are repeated here to keep the appendix self-contained. Let X and Y be totally ordered sets. A function $K(x, y)$ on $X \times Y$ is *totally positive of order 2* (TP_2) if

$$\begin{vmatrix} K(x_1, y_1) & K(x_1, y_2) \\ K(x_2, y_1) & K(x_2, y_2) \end{vmatrix} \geq 0$$

for all $x_1 \leq x_2$ in X and $y_1 \leq y_2$ in Y.

A non-negative function $h(x)$ on $R = (-\infty, \infty)$ is a *Pólya frequency function of order 2* (PF_2) if $K(x, y) = h(x - y)$ is TP_2 on $R \times R$.

$F(t)$ has a continuous density function $f(t)$ and $F(t) = \int_{-\infty}^{t} f(x)dx$. The k-fold convolution of $F(t)$ can be written in terms of $f(t)$ as follows:

$$F^{(k)}(t) = \int F^{(k-1)}(t-x)f(x)dx = \int F^{(k-1)}(u-x)f(x-v)dx,$$

where $u - v = t$. This form for the convolution will be useful later on.

Lemma A1. A non-negative function $h(x)$ is PF_2 if and only if

$$\begin{vmatrix} h(a+\delta) & h(a) \\ h(b+\delta) & h(b) \end{vmatrix} \geq 0$$

for all $a, b, \delta \in R$ satisfying $a \leq b$ and $\delta \geq 0$.

Proof: Using the equations $a = x_1 - y_2$, $b = x_2 - y_2$, and $\delta = y_2 - y_1$ to relate the variables x_1, x_2, y_1, y_2 and a, b, δ, the lemma follows. □

Let $\boldsymbol{I}(h) = \{t \mid h(t) > 0\}$.

Lemma A2. A non-negative function $h(t)$ is PF_2 if and only if the set $\boldsymbol{I}(h)$ is an interval and the ratio $\dfrac{h(t+x)}{h(t)}$ is decreasing in t for each $x \geq 0$ whenever t belongs to $\boldsymbol{I}(h)$.

Define $u(t)$ for t on the interval $\boldsymbol{I}(F)$ where

$$u(t) = \lim_{x \to 0} \frac{1}{x} \frac{F(t+x) - F(t)}{F(t)} = \frac{f(t)}{F(t)}.$$

Lemma A3. $F(t)$ is a distribution function. $F(t)$ is PF_2 if and only if $u(t)$ is decreasing for increasing t in $\boldsymbol{I}(F)$.

Proof: Since $F(t)$ is a distribution function $\boldsymbol{I}(F)$ is an interval. Assume $F(t)$ is PF_2. The idea is to look at

$$u(t) - u(t+\varepsilon) = \lim_{x \to 0} \frac{1}{x}\left[\frac{F(t+x)}{F(t)} - \frac{F(t+\varepsilon+x)}{F(t+\varepsilon)}\right]$$

for t and $t + \varepsilon$ in $\boldsymbol{I}(F)$.

The term in the braces is non-negative for every $x > 0$. This follows from Lemma A2, *i.e.*, $F(t)$ in PF_2 implies that the ratio $\frac{F(t+x)}{F(t)}$ is decreasing in t for each $x > 0$.

Conversely, assume that the ratio $u(t) = \frac{f(t)}{F(t)}$ is decreasing with increasing t in $\boldsymbol{I}(F)$. Notice that

$$\int_t^{t+x} u(s)ds = \int_t^{t+x} \frac{f(s)}{F(s)}ds = \log F(t+x) - \log F(t) = \log \frac{F(t+x)}{F(t)}.$$

Therefore

$$\frac{F(t+x)}{F(t)} = exp(\int_t^{t+x} u(s)ds).$$

It is easy to see from the above expression that if $u(s)$ is decreasing in s then the ratio $\frac{F(t+x)}{F(t)}$ is decreasing in t for each $x \geq 0$. By Lemma A2 $F(t)$ is PF_2. □

Theorem A1. If the distribution functions F_1 and F_2 are PF_2, then their convolution H, given by

$$H(t) = \int_{-\infty}^{\infty} F_1(t-s)f_2(s)ds,$$

is also PF_2.

Proof: Let F_1 and F_2 have continuous densities f_1 and f_2, respectively. Choose $x_1 \leq x_2$ and $y_1 \leq y_2$. Form

$$D = \begin{vmatrix} H(x_1 - y_1) & H(x_1 - y_2) \\ H(x_2 - y_1) & H(x_2 - y_2) \end{vmatrix}$$

where $H(x_i - y_j) = \int_{-\infty}^{\infty} F_1(x_i - s)f_2(s - y_j)ds$. Applying the generalized Cauchy-Binet identity [Karlin 1968] we get

$$D = \int\int_{s_1 < s_2} \begin{vmatrix} F_1(x_1 - s_1) & F_1(x_1 - s_2) \\ F_1(x_2 - s_1) & F_1(x_2 - s_2) \end{vmatrix} \begin{vmatrix} f_2(s_1 - y_1) & f_2(s_1 - y_2) \\ f_2(s_2 - y_1) & f_2(s_2 - y_2) \end{vmatrix} ds_1 ds_2 .$$

Integrating by parts with respect to s_1 yields

$$D = \int\int_{s_1 < s_2} \begin{vmatrix} f_1(x_1 - s_1) & F_1(x_1 - s_2) \\ f_1(x_2 - s_1) & F_1(x_2 - s_2) \end{vmatrix} \begin{vmatrix} F_2(s_1 - y_1) & F_2(s_1 - y_2) \\ f_2(s_2 - y_1) & f_2(s_2 - y_2) \end{vmatrix} ds_1 ds_2 .$$

(The integration by parts is somewhat tedious. The domain of integration requires that $s_1 < s_2$. This is achieved by integrating s_1 from $-\infty$ to s_2 and then integrating s_2 from $-\infty$ to ∞. The integration by parts with respect to s_1 involves four terms. For example, one of the terms is

$$\int\int_{s_1 < s_2} F_1(x_1 - s_1) F_1(x_2 - s_2) f_2(s_1 - y_1) f_2(s_2 - y_2) ds_1 ds_2 .$$

Performing the integration by parts on this term gives

$$\int_{s_2} F_1(x_1 - s_2) F_1(x_2 - s_2) F_2(s_2 - y_1) f_2(s_2 - y_2) ds_2$$
$$+ \int\int_{s_1 < s_2} f_1(x_1 - s_1) F_1(x_2 - s_2) F_2(s_1 - y_1) f_2(s_2 - y_2) ds_1 ds_2 .$$

The remaining three terms yield similar expressions. Luckily, the four integrals over s_2 cancel and the resulting four integrals over the domain $s_1 < s_2$ can be expressed in determinant form. The net effect on the original expression of the integration by parts with respect to s_1 is to change $F_1(x_1 - s_1)$ to $f_1(x_1 - s_1)$, $F_1(x_2 - s_1)$ to $f_1(x_2 - s_1)$, $f_2(s_1 - y_1)$ to $F_2(s_1 - y_1)$, and $f_2(s_1 - y_2)$ to $F_2(s_1 - y_2)$.)

Dividing the first determinant by $F_1(x_1 - s_1) F_1(x_2 - s_1)$ (assuming the product is non-zero) gives

$$\frac{f_1(x_1 - s_1)}{F_1(x_1 - s_1)} \frac{F_1(x_2 - s_2)}{F_1(x_2 - s_1)} - \frac{f_1(x_2 - s_1)}{F_1(x_2 - s_1)} \frac{F_1(x_1 - s_2)}{F_1(x_1 - s_1)} .$$

Since F_1 is PF_2 we have by hypothesis and Lemma A3 that

$$\frac{f_1(x_1 - s_1)}{F_1(x_1 - s_1)} \geq \frac{f_1(x_2 - s_1)}{F_1(x_2 - s_1)}$$

and by hypothesis that

$$\frac{F_1(x_2 - s_2)}{F_1(x_2 - s_1)} \geq \frac{F_1(x_1 - s_2)}{F_1(x_1 - s_1)}.$$

If $F_1(x_1 - s_1)F_1(x_2 - s_1) = 0$ then $F_1(x_1 - s_2) = 0$. This follows since F_1 is a distribution function and both $x_1 - s_1$ and $x_2 - s_1$ are less than or equal to $x_1 - s_2$. It follows that the first determinant is non-negative.

Dividing the second determinant by $F_2(s_2 - y_1)F_2(s_2 - y_2)$ (assuming the product is not zero) gives

$$\frac{F_2(s_1 - y_1)}{F_2(s_2 - y_1)} \frac{f_2(s_2 - y_2)}{F_2(s_2 - y_2)} - \frac{f_2(s_2 - y_1)}{F_2(s_2 - y_1)} \frac{F_2(s_1 - y_2)}{F_2(s_2 - y_2)}.$$

Since F_2 is PF_2 we have

$$\frac{f_2(s_2 - y_2)}{F_2(s_2 - y_2)} \geq \frac{f_2(s_2 - y_1)}{F_2(s_2 - y_1)}$$

and

$$\frac{F_2(s_1 - y_1)}{F_2(s_2 - y_1)} \geq \frac{F_2(s_1 - y_2)}{F_2(s_2 - y_2)}.$$

If $F_2(s_2 - y_1)F_2(s_2 - y_2) = 0$ then $F_2(s_1 - y_2) = 0$. Thus the second determinant is non-negative.

From the above we have that $D \geq 0$. □

Appendix B: $F^{(k)}(t)$ is TP_2 in k and t

Theorem B1 [EMP 1973, Theorem 4.9] If $F(t)$ is PF_2. Then for all $z_1 \leq z_2$

$$D(k, z_1, z_2) = F^{(k)}(z_1)F^{(k+1)}(z_2) - F^{(k)}(z_2)F^{(k+1)}(z_1) \geq 0.$$

Proof: We can write $D(k, z_1, z_2)$ as follows:

$$D(k, z_1, z_2) = \int_0^{\infty} [F^{(k)}(z_1)F^{(k)}(z_2 - \theta) - F^{(k)}(z_2)F^{(k)}(z_1 - \theta)]\, dF(\theta).$$

By setting $x_1 = z_1$, $x_2 = z_2$, $y_1 = 0$, and $y_2 = \theta$, we can apply the Theorem from the previous Appendix and conclude that the integrand is non-negative for all $\theta \geq 0$. Therefore $D(k, z_1, z_2) \geq 0$. □

From this Theorem it is easy to show that if $F(t)$ is PF_2 then $F^{(k)}(t)$ is TP_2 in k and t.

Appendix C: PF2 Densities and Distributions

In this appendix we give examples of PF_2 density and distribution functions. We provide a proof that if f is a PF_2 density, then the corresponding distribution function is PF_2. The converse does not necessarily hold as demonstrated by the examples given at the end of this appendix. In [BP 1975], Lemma 5.8, it is shown that if f is PF_2, then $\bar{F}$ is PF_2.

Lemma C1. If the density f is PF_2 then the corresponding distribution function F is PF_2.

Proof: Write $F(t) = \int_{-\infty}^{t} f(x)dx = \int_0^{\infty} f(t - x)dx$. Choose $a \leq b, \delta \geq 0$ in X and form

$$D = \begin{vmatrix} F(a + \delta) & F(a) \\ F(b + \delta) & F(b) \end{vmatrix}$$

$$= \int_0^{\infty}\int_0^{\infty} [f(a + \delta - z_2)f(b - z_1) - f(a - z_1)f(b + \delta - z_2)]\, dz_1 dz_2,$$

Let $I(z_1, z_2) = f(a + \delta - z_2)f(b - z_1) - f(a - z_1)f(b + \delta - z_2)$. It is easy to check that $I(z_1, z_2) = -I(z'_1, z'_2)$ whenever $z_2 = z'_1 + \delta$ and $z'_2 = z_1 + \delta$. This relationship establishes a 1-1 correspondence between pairs (z_1, z_2) in the region $z_1 \geq 0$ and $z_2 \geq z_1 + \delta$ and pairs (z'_1, z'_2) in the region $z'_1 \geq 0$ and $\delta \leq z'_2 \leq z'_1 + \delta$. Since $I(z_1, z_2) + I(z'_1, z'_2) = 0$ at corresponding pairs, the integral vanishes over the region consisting of $z_1 \geq 0$ and $z_2 \geq \delta$. Accordingly,

$$D = \int_0^{\delta}\int_0^{\infty} I(z_1, z_2)dz_1dz_2 .$$

Since $f(a - z_1) = 0$ for $z_1 > a$ we get

$$D \geq \int_0^{\delta}\int_0^{a} I(z_1, z_2)dz_1dz_2$$

Let $a' = a - z_1$, $b' = b - z_1$, $\delta' = (\delta - z_2) + z_1$. Since $0 \leq z_2 \leq \delta$ and $0 \leq z_1 \leq a$, it follows that $a' \leq b'$ and $\delta' \geq 0$. Since f is PF_2

$$I(z_1, z_2) = f(a' + \delta')f(b') - f(a')f(b' + \delta') \geq 0$$

for all $0 \leq z_2 \leq \delta$ and $0 \leq z_1 \leq a$. Therefore $D \geq 0$. □

The following densities are PF_2 and consequently, the corresponding distribution functions are also PF_2.

PF_2 Densities

1. Gamma: $g_{\lambda, \alpha}(t) = \dfrac{\lambda (\lambda t)^{\alpha - 1}}{\Gamma(\alpha)} e^{-\lambda t}$, $t > 0$, $\lambda > 0$, $\alpha \geq 1$.

2. Weibull: $f_{\beta}(t) = \beta\lambda (\lambda t)^{\beta - 1} e^{-(\lambda t)^{\beta}}$, $t > 0$, $\lambda > 0$, $\beta \geq 1$.

3. Exponential: Gamma with $\alpha = 1$ or Weibull with $\beta = 1$.

4. Truncated normal: $f(t) = \dfrac{e^{-(t - \mu)^2/2\sigma^2}}{a\sqrt{2\pi}\sigma}$, $t \geq 0$, $\sigma > 0$.

5. Gumbel: $g_{\beta, \lambda}(t) = \beta\lambda \exp(\lambda t - \beta(\exp(\lambda t) - 1))$, $t \geq 0$, $\beta > 0$, $\lambda > 0$.

6. Uniform: $f(t) = \text{constant}$.

7. Hypoexponential: $f_{\lambda_1, \lambda_2}(t) = \dfrac{\lambda_1\lambda_2}{\lambda_2 - \lambda_1}(e^{-\lambda_1 t} - e^{-\lambda_2 t})$ $\lambda_1 > 0$, $\lambda_2 > 0$.

Non-PF$_2$ Densities

The following densities are *not* PF$_2$.

1. Gamma with $\alpha < 1$.
2. Weibull with $\beta < 1$.
3. Pareto: $f_a(t) = at^{-(a+1)}$, $t \geq 1$ and $a > 0$.

PF$_2$ Distribution Functions

1. Gamma for all $\alpha > 0$.
2. Weibull for all $\beta > 0$.
3. Pareto for all $a > 0$.

These three cases are interesting since their densities are not PF$_2$.