DISTRIBUTIONAL CHARACTERISTICS OF THE OBSERVED CONFIDENCE COEFFICIENT IN CERTAIN RELIABILITY PROBLEMS

By

RICHARD HAROLD BROWNE

Bachelor of Science
University of Missouri at Rolla
Rolla, Missouri
1968

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1970

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
May, 1973

DISTRIBUTIONAL CHARACTERISTICS OF THE
OBSERVED CONFIDENCE COEFFICIENT
IN CERTAIN RELIABILITY
PROBLEMS

Thesis Approved:

J. Leroy Folks
Thesis Adviser

P. Larry Claypool

Lyle D. Broemeling

Hamed K. Eldin

David L. Weeks

D. D. Durham
Dean of the Graduate College

ACKNOWLEDGMENTS

This author is indebted to Dr. J. Leroy Folks for suggesting the thesis problem and for his guidance in the preparation of this thesis.

An expression of appreciation is extended to Professors David L. Weeks, Lyle D. Broemeling, P. Larry Claypool, and Hamed K. Eldin for serving on my advisory committee and for many helpful suggestions in the final draft of this thesis.

I would also like to express my sincere gratitude to those individuals outside of college who have played signal roles in my personal and professional development, culminating in the writing of this thesis.

The greatest expression of gratitude must go to my parents, Mr. and Mrs. Basil Browne, without whose lifelong support and encouragement this thesis might never have been written. Gratitude is also expressed to the following individuals for the accompanying reasons: to Thomas Moeller for sparking my interest in mathematics and encouraging me to excel in it, to Mrs. Henrietta Steinkamp for her personal encouragement and for serving as a model of what I have ever since thought a teacher should be, to Dr. Robert Davis for introducing me to the orderly nature of serious research, and to Dr. Edmund Gehan for helping to show me what the practice of statistics was really like.

Thanks are extended to John Palmer for many helpful and entertaining conversations which helped ease the burdens of thesis writing. I would also like to express my thanks to Velda Davis for her professional typing of a rather difficult thesis. Lastly, I would like to give a most special thanks to my wife, for whom no words are equal to her contribution.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

R	actual reliability
β	desired reliability
QL	qualification limit
γ	observed confidence coefficient
c.d.f.	cumulative distribution function
d.f.	degrees of freedom
SL	observed significance level
$D(\gamma)$	distribution function of γ
$d(\gamma)$	density function of γ
OSTL	one-sided tolerance limit
S	a random sample of size n
$T_d(\cdot)$	c.d.f. of a noncentral-t with parameter d and n-1 d.f.
$t_d(\cdot)$	density function of a noncentral-t with parameter d and n-1 d.f.
$G_n(\cdot)$	c.d.f. of a gamma with parameter n
$g_n(\cdot)$	density of a gamma with parameter n
$N(\mu,\sigma^2)$	a normal distribution with mean μ and variance σ^2
$IN(\beta)$	the point below which β of a N(0,1) distribution lies
$CN_d(\cdot)$	c.d.f. of a N(d,1) random variable
L(S)	value of a lower tolerance limit for a sample S

CHAPTER I

INTRODUCTION

Background

While the topic of reliability assessment did not appear in the literature prior to the Second World War, it is a relatively old concept. Generally speaking, the reliability of a product is the probability that it will meet or exceed certain predetermined criteria. For example, the reliability of a light bulb might be the probability that it will burn for at least 500 hours. However, since all light bulbs produced by one manufacturer are not exactly alike, this probability could never be known with certainty unless all the bulbs were tested. For this reason, statistical methods known as reliability assessment methods were developed to help quantify the uncertainty about the true reliability of the product.

Reliability assessment had its beginnings in the methods of quality control. Generally speaking, quality control is the practice of accepting or rejecting lots of a product based on the performance of samples drawn from those lots. Since it is usually unreasonable to expect every item in the lot to meet the performance criteria, an agreement is reached beforehand on an acceptable fraction of unsatisfactory items which may be in a lot. The manufacturere and buyer must also decide what chance they would want to take of discarding a good lot or buying a bad lot, respectively, based on a sample from that lot. Using statistical methods, standards are set for the sample which will determine whether the lot is to be accepted or rejected. The desired fraction of acceptable items in a lot could be referred to as the reliability of an item drawn at random from the lot since it gives the probability of an item meeting the criteria in a sampling frequency sense. However, the term "reliability" did not appear in the quality control literature prior to the 1950's.

Literature on quality control was available prior to the Second World War (e.g., (1) (2) (3)), primarily of the nonparametric variety. The outbreak of hostilities with its vastly increased production of war materials, though, was the major catalyst to the development of more and better techniques of quality control (e.g., (4) (5) (6)).

One quality control method that was more fully developed during the war was a method called tolerance limits. Briefly, tolerance limits are values computed from data on a random sample of items between which (say) 90% of the product lifetimes fall with (say) 99% confidence. (The meaning of "confidence" will be discussed below.) The fraction of lifetimes which falls between the tolerance limits can be thought of as the reliability associated with the operating criteria given by the tolerance limits. For example, suppose the tolerance limits on 90% of light bulb lifetimes are computed to be from 52 hours to 639 hours, with 99% confidence. This means that one can be "99% confident" that the light bulbs are at least 90% reliable with respect to the operating criteria of lasting more than 52 hours but less than 639 hours. Stated another way, one would be "99% confident" that one could pick a light bulb at random 9 out of 10 times that would burn between 52 hours and 639 hours. Thus, tolerance limits provide a more direct means of making a statement about the reliability of a product than earlier methods in quality control.

Much of the work done on tolerance limits during the war was by Wald, Wilks, and Wolfowitz (e.g., (7) (8) (9) (10) (11). While much of the work was of a nonparametric nature, some work was done with the assumption of a normal distribution for the product lifetimes (7) (8). As understanding of engineering systems increased, tolerance limits were developed for such lifetime distributions as the exponential (12) (13), the Weibull (13), and the gamma (13). Much of the literature of the 1950's on tolerance limits, life testing, and reliability is given in an article by Mendenhall (14).

Tolerance limits are also used to find the confidence with which one can say that a product is (say) 95% reliable (15). Exactly how this confidence is computed will be discussed in a later chapter. Generally, one first establishes the specification limit or qualification limit (QL) that the product should meet or exceed. For example, one may want his light bulbs to burn at least 500 hours. Next, a sample of products is tested. One then finds the confidence γ such that if one used that confidence and the sample data, one would compute a tolerance limit equal to the QL. This computed confidence γ and its usage are the primary subjects of this dissertation.

This computed confidence γ is more commonly referred to as the observed confidence or observed confidence coefficient. Little has been written about the observed confidence; however, much has been written about the observed significance level (SL), which is related to γ by the identity

$$\gamma = 1 - SL$$

or $$SL = 1 - \gamma \ .$$

It is apparent that any discussion about the nature of the SL would easily carry over to that for γ (and vice-versa).

A Review of the Observed Significance Level

For a more complete discussion of the SL, one is referred to the text by Kempthorne and Folks (16) and papers by Anscombe (17) and Fisher (18). The following is a brief summary of the use and meaning of the SL in experimentation. While hardly rigorous, this presentation will serve to illustrate the uses and abuses of the SL and, concomitantly, γ.

Before beginning experimentation, one states some hypotheses about the true state of nature, such as the reliability of G.E. light bulbs, height of the average O.S.U. student, etc. This statement is commonly referred to as the null hypothesis (H) about the population of interest. To help determine the validity of H, one draws a random sample from the population of interest. One uses this sample to determine the degree of support (or validity) for H provided by the data. If H was that the average student height at O.S.U. was 72 inches and the average height of 100 randomly chosen students was only 64 inches, this result would cast much doubt upon the validity of H. Conversely, if the average was 71.5 inches, one could easily believe that the 100 students could have been drawn from a population whose overall average height was 72 inches. The SL provides one way of quantifying the degree of support given H by the data.

The SL is computed as the probability of drawing a sample from the population no more in agreement with H than the sample actually drawn.

If SL is near zero, the sample obtained would be extremely rare if H is true. Under those circumstances, one might begin looking for another H which might better explain the observed data. Conversely, if SL is substantially greater than zero, this would say that the observed sample would not be unreasonable to expect if H were true. In that case, one would probably be satisfied to retain H as the apparent true state of nature.

Among the many misuses of the SL is the interpretation of the SL as the probability that H is true. Obviously, H is true or false, irrespective of the data. Hence, the probability that H is true is either zero or one, not equal to SL. Another misuse stems from the notion of accepting or rejecting H, a practice taught in most statistical methods books (e.g., (19) (20)). If, prior to experimentation, one states that he will accept H if the SL $> p_o$ (a predetermined value) and reject H otherwise, then he will reject a true H in nearly $100p_o$% of his experiments. The value p_o is commonly referred to as the probability of a Type I error. Clearly, this scheme is of value in lot inspection problems, where p_o is the chance of destroying a good lot using this method.

However, a common erroneous usage of the SL is to look at the SL from one experiment as the risk or chance of making an error if H is rejected when in fact it is true. For example, in testing missile components, one sample of components might yield an SL = 0.10. The conclusion usually drawn is that one would run a 10% risk of being wrong if he rejected this lot of missile components. The fallacy in this thinking is apparent when one realizes that the SL is a function of the data and so will be different for each sample drawn from the <u>same</u> lot.

A manufacturer who would choose to keep a lot with only a "10% chance of being in error" might be surprised to find the risk computed from another sample from the lot to be more than enough to make him reject it (say 1%). Since the SL is strictly a function of the sample, it is hard to see how it could really be used as probability of error.

The correct use of the SL from a single experiment is as a measure of the support given H by the data. How one chooses to weigh the SL in his evaluation is strictly a subjective matter; experience is usually the best gauge. Attempts to use values of the SL below 0.05 or 0.01 in deciding whether to accept or reject H are quite common. Such practices can usually be countered by asking what an SL of 0.05 or 0.01 means in the experimental context. Usually there is little basis for the use of such levels other than tradition.

To compute the observed confidence γ, one need only compute the SL and then $\gamma = 1 - \text{SL}$. While not normally stated as such, γ gives the degree of support for H provided by the data. If γ is near unity, this is support for H; if γ is substantially less than unity, this is evidence against H.

Clearly, since the SL is a function of the data (i.e., a statistic), so is γ. As a statistic, γ has distributional characteristics which would be affected by the sample size, H, and the true state of nature. What these characteristics are and how they might influence the use of γ by experimenters is the subject of the following chapters.

Selection of Problems

The problems in this dissertation refer to the choice of lifetime distributions and the tolerance limit formulas. The lifetime distributions will be the normal (variance known and unknown), gamma, Weibull, and exponential. The tolerance limits will be primarily of the lower limit type. Two-sided tolerance limits for a normal distribution will also be considered.

Two other distributions of interest to reliability assessers will also be considered. These will be the binomial and the Poisson. These distributions count the number of items on trial which fail to meet the desired QL.

Since the results for the confidence coefficient γ carry over to significance testing via the SL, a chapter will consider the impact of these findings on significance testing. Since little has been done in the way of examining particular significance tests for their distributional characteristics, this could prove to be a fruitful area of research.

CHAPTER II

EXPRESSING THE CONFIDENCE COEFFICIENT AS A STATISTIC

A matter of frequent interest to manufacturers is knowing the proportion of their products which exceeds certain design criteria. Unfortunately, it is often impossible to know this proportion with certainty since the operational parameters (e.g., failure rate) are unknown. By testing a random sample of products, though, one can use the test data to compute

(a) a limit on performance which at least (say) 95% of the products will meet with (say) 99% confidence, or

(b) the confidence with which one can say that 95% of the products meet a predetermined level of performance.

This first problem is called finding a tolerance limit while the second is called finding the observed confidence coefficient. How these two problems are interrelated is illustrated in the following discussion.

One-Sided Tolerance Limits

In general, the one-sided tolerance limit (OSTL) is an expression, say TL(S), that, for a sample S, has probability γ of providing a value above which lies a proportion β of the population. For a particular sample S_o, the value $TL(S_o)$ has γ confidence of bounding β of the population. In reliability problems, $100\beta\%$ is usually the desired reliability. The two-sided tolerance limits, on the other hand, use two formulas to bound β of the population. The topic of two-sided tolerance limits will be taken up in a later chapter.

The OSTL problem has two forms:

(a) a lower OSTL on at least β of the population,

(b) an upper OSTL on at most β of the population.

It will be shown that the other two possible forms

(c) an upper OSTL on at most $1-\beta$ of the population,

(d) a lower OSTL on at least $1-\beta$ of the population

are actually the same as (a) and (b), respectively.

Let S be a random sample of size n from a population with density function $f(\cdot)$. In a reliability problem, $f(\cdot)$ could be the density of lifetimes for the products. Then L(S) is a lower OSTL on at least β of the population with probability γ if

$$\Pr\left(\int_{L(S)}^{\infty} f(y)dy \geq \beta\right) = \gamma . \tag{1}$$

Note that this is the same as

$$\Pr\left(1 - \int_{-\infty}^{L(S)} f(y)dy \geq \beta\right) = \gamma$$

or

$$\Pr\left(\int_{-\infty}^{L(S)} f(y)dy \leq 1-\beta\right) = \gamma . \tag{2}$$

Comparing (1) and (2), it is apparent that L(S) can perform the functions stated in (a) and (c).

In similar fashion, U(S) is an upper OSTL on at most of the population if

$$\Pr\left(\int_{-\infty}^{U(S)} f(y)\,dy < \beta\right) = \gamma \, . \tag{3}$$

This is also expressible as

$$\Pr\left(\int_{U(S)}^{\infty} f(y)\,dy \geq 1-\beta\right) = \gamma \, . \tag{4}$$

Comparing (3) and (4), it is apparent that $U(S)$ can perform the functions stated in (b) and (d), respectively.

The following discussion will be primarily centered about the OSTL in (a) (and (c)), since this is the form of most interest to people in life testing. Results for the (b) and (d) cases could be obtained by slight modifications of the techniques used below.

The Confidence Coefficient in a Reliability Problem

Suppose a product must function for at least B hours to be considered acceptable. The user wishes to know with what confidence he can say that his products are at least 100β% reliable, based on the results of a sample. That is, how confident can he be that at least β of his production will have lifetimes greater than B?

In the OSTL problem, one asks: given β, γ, and a sample S_o, what must $L(S_o)$ be so that one can be γ-confident that $L(S_o)$ is a lower bound on at least β of the lifetimes? To find the observed confidence γ, though, one asks: given β, B, and S_o, what must γ be so that $L(S_o)$ would be equal to B? The value of γ is the observed confidence for that sample S_o. Surely, as S_o changes as different samples are taken, the value of γ must change so that $L(S_o)$ will always equal B. In this way, the observed confidence γ is a function of the data (i.e., a statistic). As such, it has a distribution depending upon the type of lifetime population, the tolerance limit formula, B, n, and β. In the next chapter, the population parameters and B will be reparameterized into the parameter R, the true fraction of the population lying to the right of B. This will allow for a more general tabulation of the results for the confidence coefficient.

CHAPTER III

RESULTS FOR SOME CONTINUOUS LIFETIME DISTRIBUTIONS

For each lifetime distribution considered below, one OSTL formula was used to find the confidence coefficient γ. While there is certainly more than one OSTL for each population, the OSTL's chosen here appear to have "good" properties and have appeared in reputable journals. The populations considered appear to be those that appear most commonly in reliability problems.

Normal Distribution With Variance Unknown

Let S be a random sample of size n from a normally distributed population with unknown parameters μ and σ^2. Let $L(S) = \bar{x} - ks$ be a γ-probability lower OSTL on at least β of the population, where k is a constant to be determined. Then

$$\Pr\left(\int_{\bar{x}-ks}^{\infty} dN(\mu,\sigma^2) \geq \beta\right) = \gamma$$

where $N(\mu,\sigma^2)$ is the distribution function of a normal random variable with mean μ and variance σ^2. The above equation can also be written as

$$\Pr\left(1 - \int_{-\infty}^{\bar{x}-ks} dN(\mu,\sigma^2)\right) \geq \beta) = \gamma$$

$$= \Pr\left(\int_{-\infty}^{\frac{\bar{x}-ks-\mu}{\sigma}} dN(0,1) \leq 1 - \beta\right)$$

$$= \Pr((\bar{x} - ks - \mu)/\sigma \leq IN(1 - \beta))$$

where IN(p) is the point below which the fraction p of a standard normal distribution lies. So

$$\gamma = \Pr(\sqrt{n}\left(\frac{\bar{x} - \mu}{\sigma} - IN(1 - \beta)\right) \leq \sqrt{n}ks/\sigma)$$

$$= \Pr\left\{\sqrt{n}\left[\frac{\frac{\bar{x} - \mu}{\sigma} - IN(1 - \beta)}{S/\sigma}\right] \leq \sqrt{n}\ k\right\}.$$

But the term on the left is a noncentral-t random variable with parameter $\delta = \sqrt{n}IN(1 - \beta) = \sqrt{n}IN(\beta)$. Then

$$\gamma = T_\delta(\sqrt{n}k)$$

where $T_\delta(\cdot)$ is the distribution function of the noncentral-t random variable with parameter δ and n-1 degrees of freedom (d.f.). Solving the above equation for k yields

$$k = T_\delta^{-1}(\gamma)/\sqrt{n}$$

and so L(S) can be written as

$$L(S) = \bar{x} - T_\delta^{-1}(\gamma)s/\sqrt{n}.$$

Suppose a manufacturer establishes a minimum WL of B hours of operation. The value of γ required so that L(S) will equal B is found by solving

$$B = \bar{x} - T_\delta^{-1}(\gamma)s/\sqrt{n}$$

which yields

$$\gamma = T_\delta(\sqrt{n}(\bar{x} - B)/s). \tag{1}$$

Clearly, this establishes γ as a function of the data. This writer acknowledges the 1968 article by Owen (19) which provided the inspiration for this approach to finding γ. The following is an example of how Equation (1) might be used to find the confidence in a reliability assessment problem. Suppose a tire manufacturer wants to know with what confidence he can claim that at least 95% of his tires will last at least 12,000 miles. He puts 9 tires on test and finds that the average tire life was 18,400 miles with a standard deviation for the 9 tires of 3,000 miles. To use (1), of course, one must assume that tire life is normally distributed. Assuming this, the value of the noncentrality parameter $\delta = \sqrt{9}\text{IN}(.95) = 3(1.645) = 3.935$. The value of $\sqrt{n}(\bar{x} - B)/s = 6.4$. Then $\gamma = T_{3.935}(6.4)$. To find this value, the tables by Resnikoff and Liebermann (20) on the noncentral-t were used. With a little interpolation, the value of γ is found to be approximately 0.80. Thus, the manufacturer can claim his tires are 95% reliable with about 80% confidence.

Since γ is a statistic, it will have a distribution function. The distribution function of γ will be denoted by $D(\gamma)$ in this and all other developments. In this case,

$$D(\gamma) = \Pr(T_\delta(\sqrt{n}(\bar{x} - b)/s) \leq \gamma), \quad 0 < \gamma < 1$$

$$= \Pr(\sqrt{n}(\bar{x} - B)/s \leq T_\delta^{-1}(\gamma)).$$

But $\sqrt{n}(\bar{x} - B)/s$ is a noncentral-t random variable with parameter $\theta = \sqrt{n}(\mu - B)/\sigma$ and $n - 1$ d.f. Thus,

$$D(\gamma) = T_\theta(T_\delta^{-1}(\gamma)), \quad 0 < \gamma < 1. \tag{2}$$

Upon rexamining θ, it is apparent that

$$-\theta = \sqrt{n}(B - \mu)/\sigma = \sqrt{n}\,\mathrm{IN}(1 - R)$$

where R is the true fraction of the distribution to the left of B. But since

$$\mathrm{IN}(1 - R) = -\mathrm{IN}(R) \quad 0 < R < 1,$$

then $\theta = \sqrt{n}\,\mathrm{IN}(R)$. Thus $D(\gamma)$ in (2) has parameters of n (sample size in test), β (desired reliability), and R (true reliability).

The density of γ is found by differentiating (2) with respect to γ. The density or probability function of γ in this and subsequent developments will be denoted by $d(\gamma)$. In this case, the density is

$$d(\gamma) = t_\theta(T_\delta^{-1}(\gamma)) \frac{dT_\delta^{-1}(\gamma)}{d\gamma}$$

where $t_\theta(\cdot)$ is the density of a noncentral-t random variable with parameter θ and n-1 d.f. It is shown in the Appendix that

$$\frac{dT_\delta^{-1}(\gamma)}{d\gamma} = \frac{1}{t_\delta(T_\delta^{-1}(\gamma))}.$$

Hence,

$$d(\gamma) = \left.\frac{t_\theta(x)}{t_\delta(x)}\right|_{x = T_\delta^{-1}(\gamma)}, \quad 0 < \gamma < 1 \tag{3}$$

According to Resnikoff and Liebermann (20), one form of the noncentral-t density with parameter δ and n-1 d.f. is

$$t_\delta(x) = \frac{f!}{2^{\frac{f-1}{2}}\Gamma(f/2)\sqrt{\pi f}}\, e^{-\frac{1}{2}\left(\frac{f\delta^2}{f+x^2}\right)}\left(\frac{f}{f+x^2}\right)^{\frac{f+1}{2}} H_f\left[\frac{-\delta x}{\sqrt{f+x^2}}\right]$$

where

$$f = n-1$$

$$H_f(y) = \int_0^\infty \frac{v^f}{f!}\, e^{-(v+y)^2/2}dv.$$

Substituting this into (3) yields

$$d(\gamma) = \exp\left[\frac{(\delta^2-\theta^2)f}{2(f+x^2)}\right]\frac{H_f\left[-\delta x/\sqrt{f+x^2}\right]}{H_f\left[-\theta x/\sqrt{f+x^2}\right]}\Bigg|_{x = T^{-1}(\gamma)}. \qquad (4)$$

If $d(\gamma)$ is monotone increasing (decreasing) then increasing values of γ are more (less) likely. This can be shown by differentiating $d(\gamma)$ and determining under what circumstances $d'(\gamma)$ is greater than or less than zero. All attempts to do this for (4) have failed, although the plots of $d(\gamma)$ in Figures 1 to 6 do suggest that $d(\gamma)$ is increasing with γ when $R > \beta$ and decreasing with γ when $R < \beta$. This says that if $R > \beta$ (the true reliability is greater than the desired reliability) that values of γ near 1 are more likely. Conversely, if $R < \beta$, values of γ near zero are more likely than values of γ greater than zero.

The values of d(0) and d(1) are not discernible from the plots. To find the values, certain limits must be considered.

$$\lim_{\gamma\to\infty} T_\delta^{-1}(\gamma) = -\infty \text{ and } \lim_{\gamma\to 1} T_\delta^{-1}(\gamma) = \infty.$$

Thus, the first term in (4) has a limiting value of 1 whether $\gamma \to 0$ or

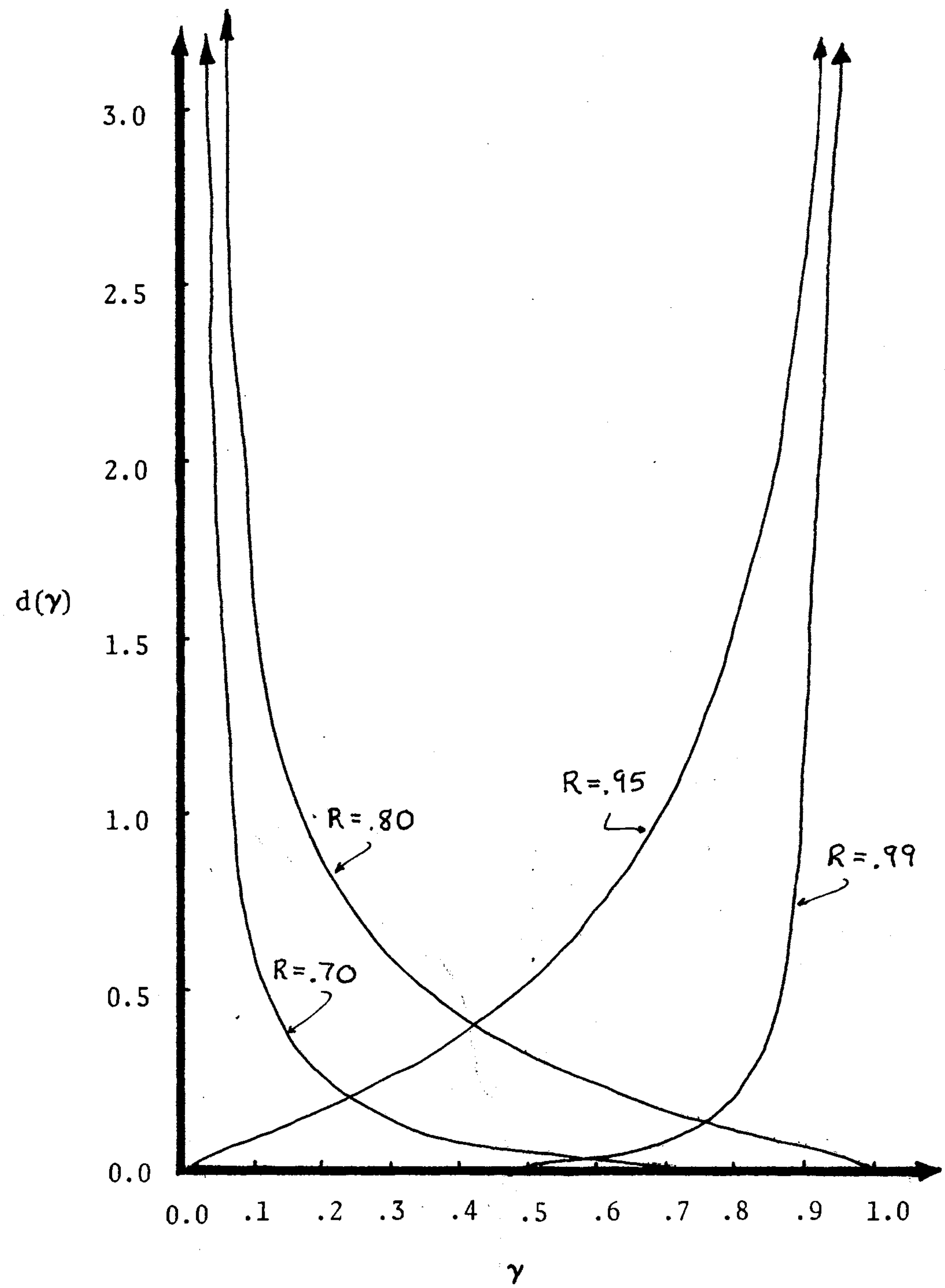


Figure 1. Densities of γ; Normal With σ Unknown; β = .90, n = 20

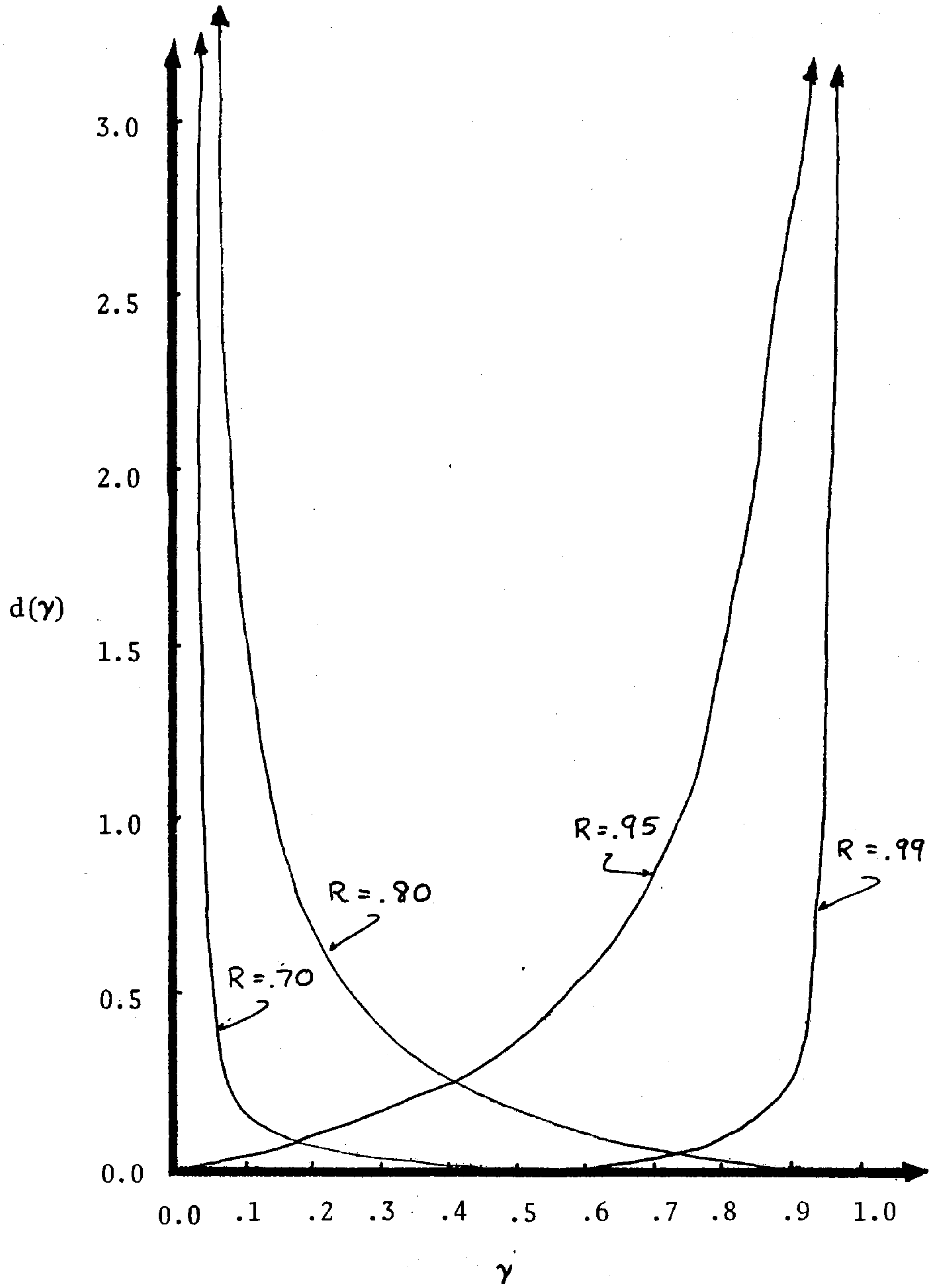


Figure 2. Densities of γ; Normal With σ Unknown;
$\beta = .90$, $n = 30$

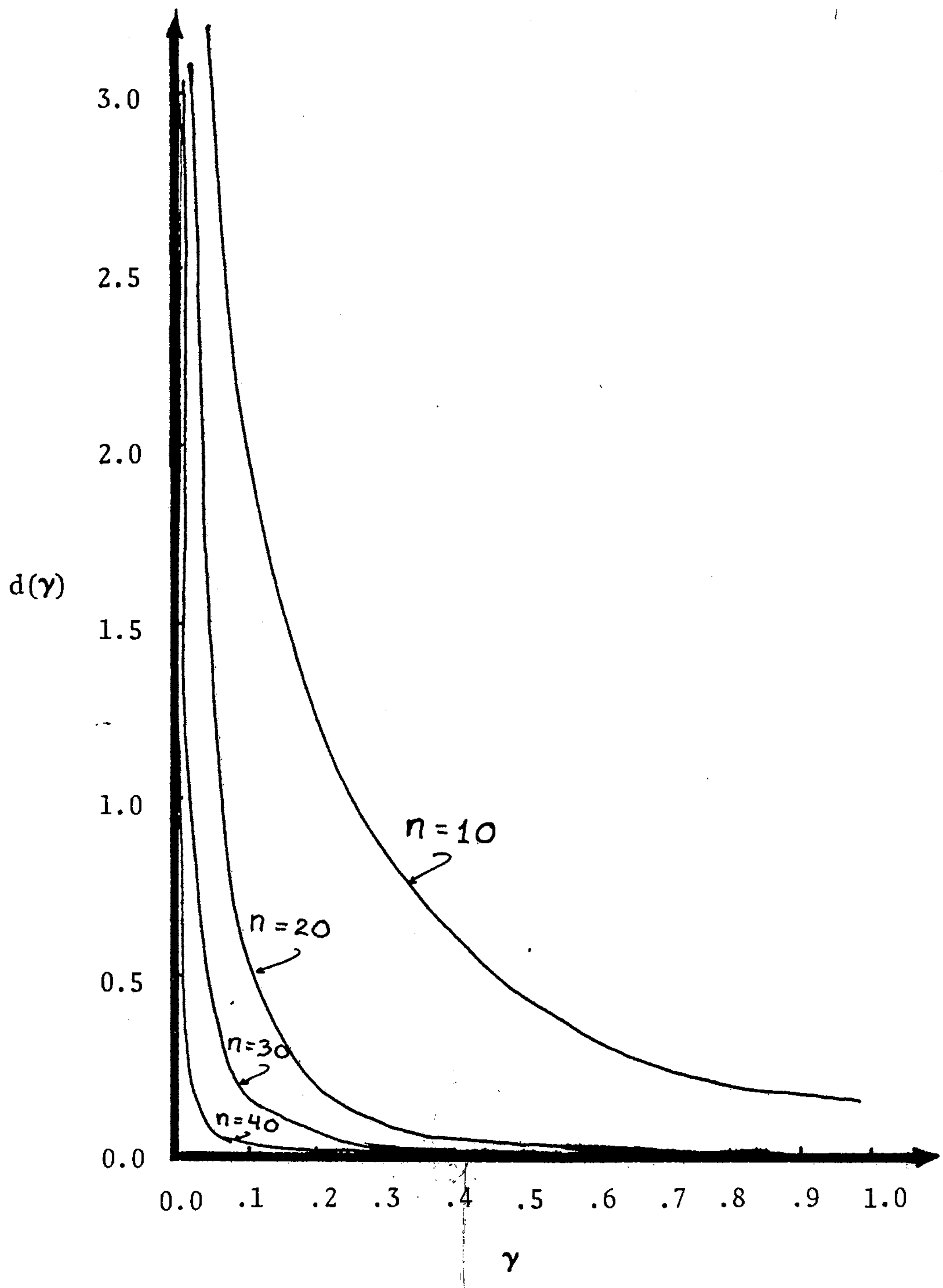


Figure 3. Densities of γ; Normal With σ Unknown; $\beta = .90$, $R = .70$

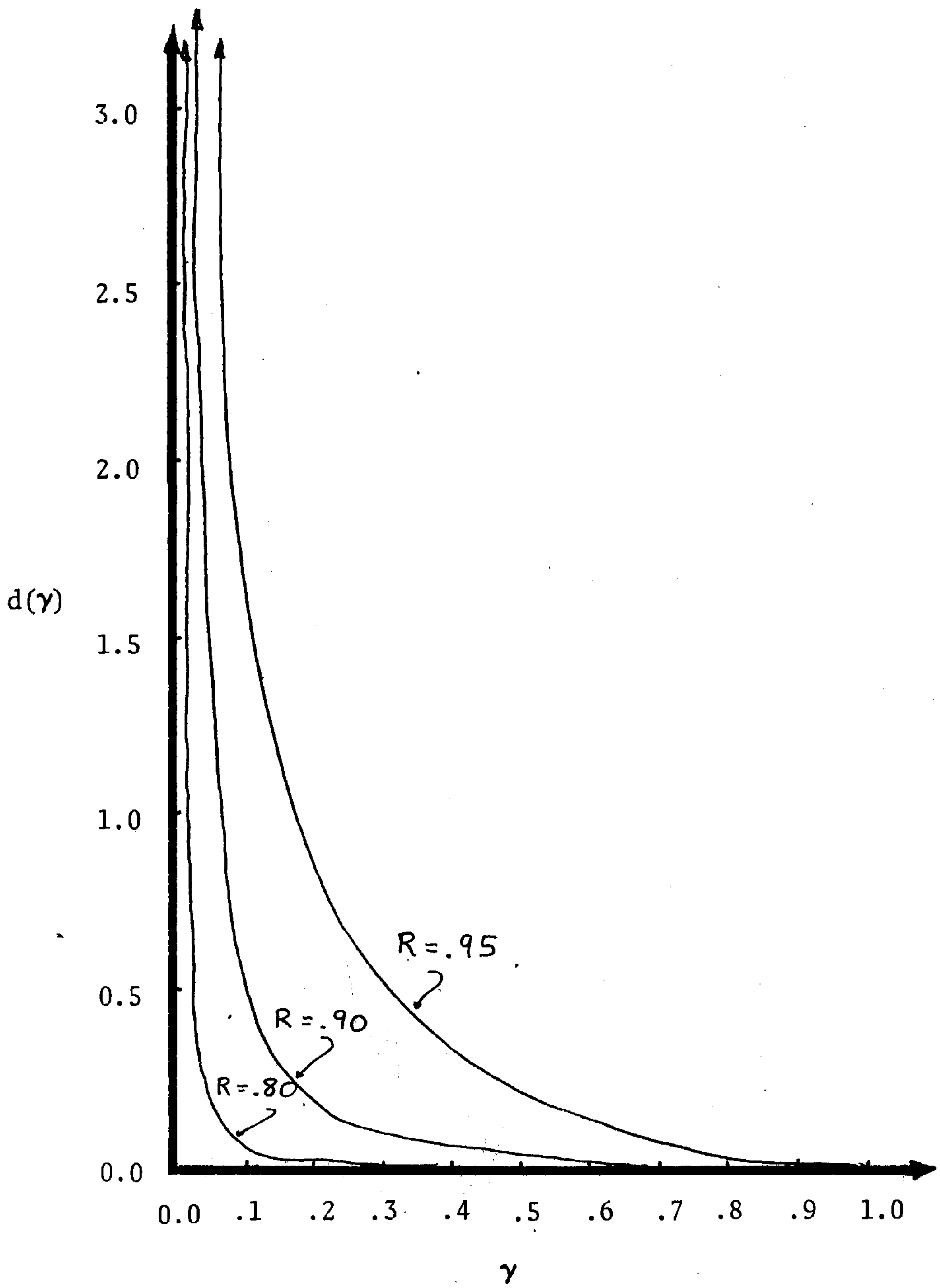


Figure 4. Densities of γ; Normal With σ Unknown; $\beta = .99$, $n = 20$

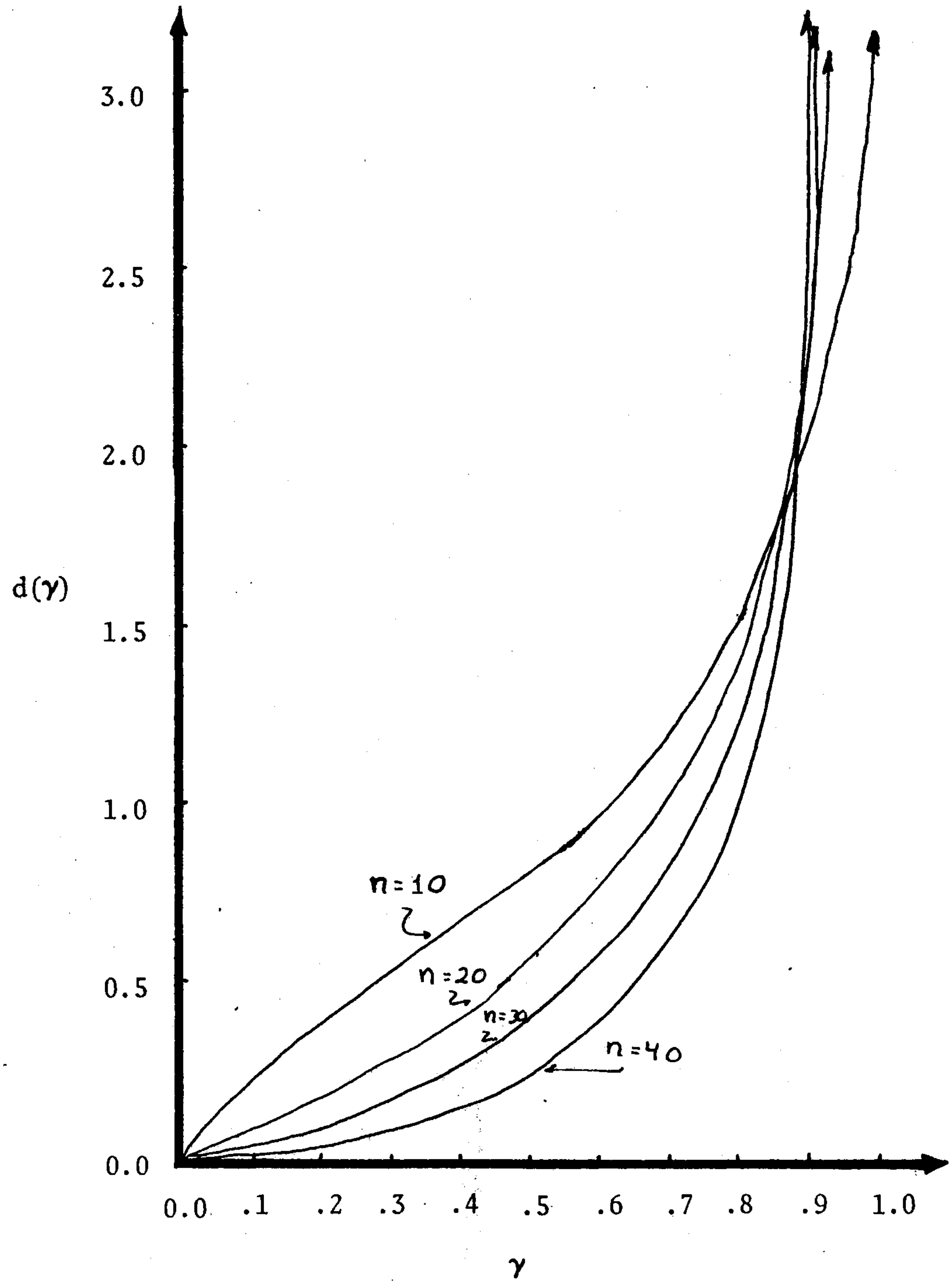


Figure 5. Densities of Y; Normal With σ Unknown; β = .95, R = .90

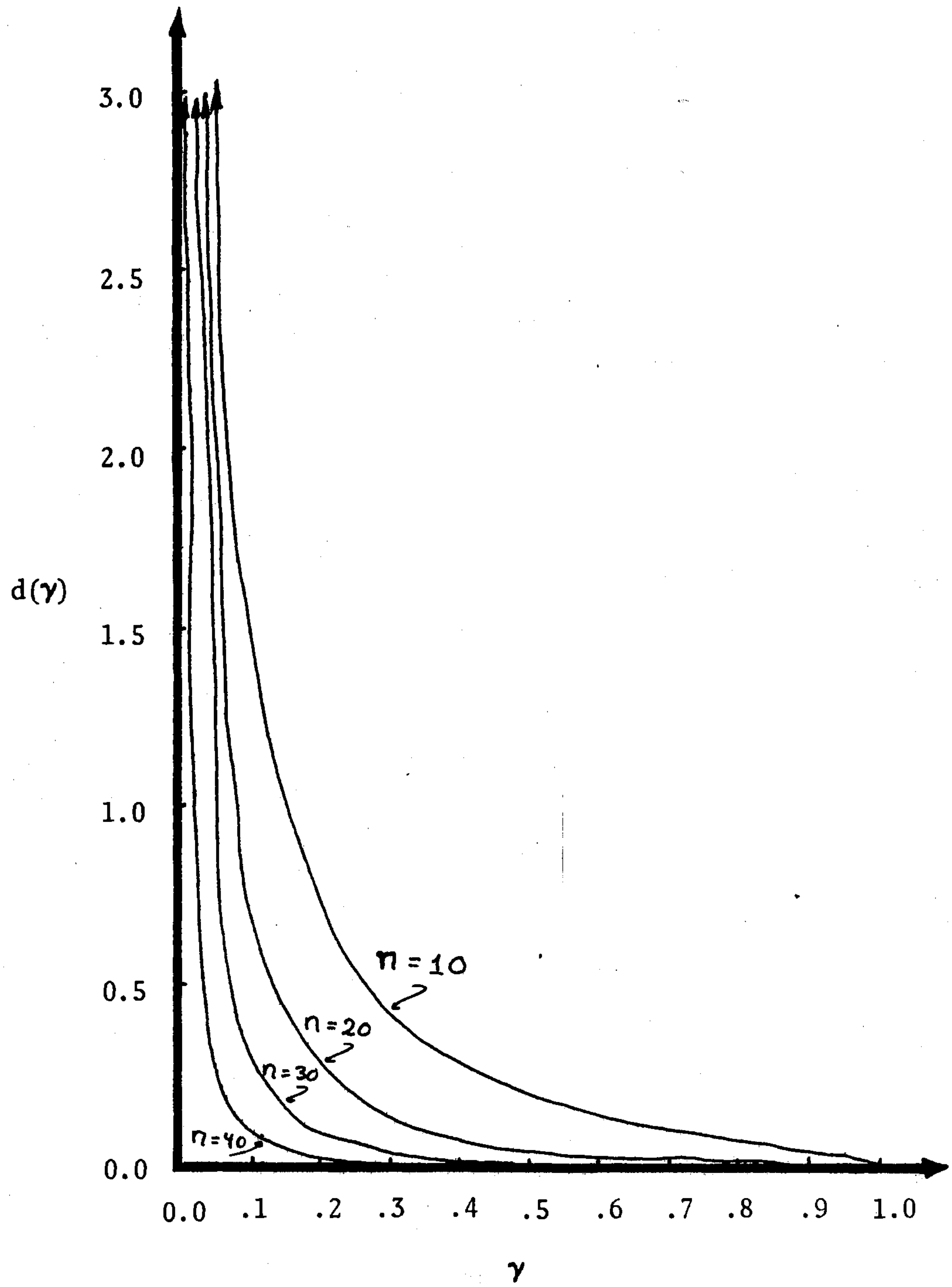


Figure 6. Densities of γ; Normal With σ Unknown; $\beta = .95$, $R = .80$

$\gamma \to 1$ since the $T_\delta^{-1}(\gamma)$ is squared. However.

$$\lim_{x \to \infty} \frac{H_f\left[-\delta x/\sqrt{f + x^2}\right]}{H_f\left[-\theta x/\sqrt{f + x^2}\right]} = \frac{H_f(-\delta)}{H_f(-\theta)}$$

while

$$\lim_{x \to \infty} \frac{H_f\left[-\delta x/\sqrt{f + x^2}\right]}{H_f\left[-\theta x/\sqrt{f + x^2}\right]} = \frac{H_f(\delta)}{H_f(\theta)} .$$

Thus, the limiting values of d(0) and d(1) are

$$d(0) = H_f(\delta)/H_f(\theta)$$

and

$$d(1) = H_f(-\delta)/H_f(-\theta).$$

The limits of $\gamma = 0$ and $\gamma = 1$ are finite since $H_f(\cdot)$ is a finite function (20).

The plots of $d(\gamma)$ seem to suggest a greater skewness towards $\gamma = 1$ when $R > \beta$ than towards $\gamma = 0$ when $R < \beta$ by an equivalent amount. This skewness becomes more exaggerated with increasing sample size. This result may in part be due to the fact that all the values of β are near 1.00. The skew is also greater the further R is from β. These early results suggest that if the true reliability (R) exceeds the desired reliability (β), a value of γ near 1.00 would be expected. Also, a value of γ near 1.00 appears unlikely if R is less than β. Such a result appears more likely with an increased sample size, which seems reasonable since one can better assess the reliability with a larger sample. If the sample size is relatively large (say 30 or more) a value of γ greater than about .95 should strongly suggest that $R > \beta$, while a value of γ below about .80 should be strong evidence that R is less than β. Values of γ in between .80 and .95 would appear to be relatively inconclusive. However, the larger the sample size, the smaller the inconclusive range would appear to be; also a higher value of γ (say .95) would be required to feel safe in concluding that $R > \beta$.

One measure of how skewed the distribution is the moments of a random variable. In order to tabulate the mean and variance of γ for some combinations of n, R, and β the following expression for the p^{th} moment of γ was developed. The p^{th} moment is given by

$$E(\gamma^p) = \int_0^1 \gamma^p d(\gamma) d\gamma$$

or

$$E(\gamma^p) = \int_0^1 \gamma^p t_\theta \left[T_\delta^{-1}(\gamma)\right] \Big/ t_\delta \left[T_\delta^{-1}(\gamma)\right] d\gamma.$$

Let

$$x = T_\delta^{-1}(\gamma)$$

$$\text{or} \quad \gamma = T_\delta(x)$$

$$d\gamma = t_\delta(x)dx.$$

Then

$$E(\gamma^p) = \int_{-\infty}^{\infty} T_\delta^p(x) t_\theta(x) dx.$$

This expression provides a non-closed form for finding $E(\gamma)$ and $Var(\gamma) = E(\gamma^2) - E^2(\gamma)$. Table I gives values of $E(\gamma)$ and $Var(\gamma)$ for selected values of n, R, and β. These calculations were done using Simpson's Rule in double precision. A method for calculating $H_f(\cdot)$ is given in (20).

TABLE I

TABULATIONS OF $E(\gamma)$ FOR THE NORMAL WITH $VAR(\gamma)$ IN PARENTHESES

n = 10	Actual Reliability (R)				
Desired Reliability (β)	.70	.80	.90	.95	.99
.90	.0829 (.0245)	.2225 (.0610)	.5000 (.0833)	.7061 (.0577)	.8965 (.1000)
.95	.0280 (.0088)	.0970 (.0392)	.2935 (.1603)	.5000 (.0833)	.7947 (.6717)
.99	.0037 (.0010)	.0181 (.0056)	.0854 (.0342)	.2005 (.0982)	.5000 (.0833)

n = 20	Actual Reliability (R)				
Desired Reliability (β)	.70	.80	.90	.95	.99
.90	.0235 (.0054)	.1365 (.0376)	.5000 (.0833)	.7794 (.0466)	.9076 (.0629)
.95	.0030 (.0005)	.0308 (.0087)	.2167 (.1046)	.5000 (.0833)	.8642 (.0742)
.99	.0001 (.0001)	.0012 (.0002)	.0236 (.0065)	.1115 (.0442)	.5000 (.0833)

n = 30	Actual Reliability (R)				
Desired Reliability (β)	.70	.80	.90	.95	.99
.90	.0073 (.0012)	.0886 (.0234)	.5000 (.0833)	.8260 (.0388)	.9134 (.0274)
.95	.0004 (.0001)	.0106 (.0022)	.1671 (.0728)	.5000 (.0833)	.9053 (.0421)
.99	.0001 (.0001)	.0001 (.0001)	.0072 (.0014)	.0662 (.0220)	.5000 (.0833)

TABLE I (Continued)

n = 40	Actual Reliability (R)				
Desired Reliability (β)	.70	.80	.90	.95	.99
.90	.0023 (.0003)	.0592 (.0145)	.5000 (.0833)	.8595 (.0329)	.9999 (.0001)
.95	.0001 (.0001)	.0038 (.0006)	.1316 (.0523)	.5000 (.0833)	.9227 (.0342)
.99	.0001 (.0001)	.0001 (.0001)	.0023 (.0003)	.0405 (.0114)	.5000 (.0833)

These limited results tend to corroborate the earlier conclusions. The ability of γ to discriminate between $R > \beta$ and $R < \beta$ appears to be quite good for large n (say 40 or more). Certainly, additional work here would appear to be potentially very fruitful.

Normal Distribution With Variance Known

Let S be a random sample of size n from a normally distributed population with unknown mean μ and known variance σ^2. Let $L(S) = \bar{x} - k\sigma$ be a γ-probability lower OSTL on at least β of the population, where k is a constant to be determined. Then

$$\gamma = \Pr\left(\int_{\bar{x}-k\sigma}^{\infty} dN(\mu,\sigma^2) \geq \beta\right)$$

$$= \Pr\left(\int_{-\infty}^{(\bar{x}-k\sigma,\mu)/\sigma} dN(0,1) \leq 1 - \beta\right)$$

$$= \Pr((\bar{x} - \mu)/\sigma - k \leq IN(1 - \beta))$$

$$= \Pr(\sqrt{n}(\bar{x} - \mu)/\sigma \leq \sqrt{n}(k + IN(1 - \beta))$$

$$= CN_o(\sqrt{n}(k + IN(1 - \beta)))$$

where

$CN_d(\cdot)$ is the cumulative distribution function of a $N(d,1)$.

Solving for k,

$$k = IN(\gamma)/\sqrt{n} - IN(1 - \beta)$$

$$\doteq IN(\gamma)/\sqrt{n} + IN(\beta).$$

Then $$L(S) = \bar{x} - (IN(\gamma)/\sqrt{n} + IN(\beta))\sigma.$$

Suppose a lower QL of B is established by a manufacturer. Then the value of the observed confidence γ is found by solving $L(S_o) = B$ in the form

$$B = \bar{x} - (IN(\gamma)/\sqrt{n} + IN(\beta)).$$

Solving this for γ yields

$$\gamma = CN_o(\sqrt{n}((\bar{x} - B)/\sigma - IN(\beta))).$$

This provides a formula for the observed confidence for a desired reliability β, sample of size n, and QL = B.

In the previous example having to do with tire reliability, assume that all conditions remain the same except that the standard deviation is known to be 3,000 miles. $IN(.95) = 1.645$ and $\sqrt{n}((\bar{x} - B)/\sigma - IN(\beta)))$ $= 2.465$. Using the cumulative normal tables in reference (16), the observed confidence γ is found to be 0.9931. This says that if the manufacturer was certain that the standard deviation of tire lifetimes is 3,000 miles, he could claim his tires are at least 95% reliable with over 99% confidence, instead of 80% when σ must be estimated from the data.

The distribution function of γ is

$$D(\gamma) = Pr(CN_o(\sqrt{n}((\bar{x} - B)/\sigma - IN(\beta))) \leq \gamma)$$

$$= Pr(\sqrt{n}(\bar{x} - B)/\sigma \leq \sqrt{n}IN(\beta) + IN(\gamma)).$$

But $\sqrt{n}(\bar{x} - B)/\sigma$ is a normal random variable with mean $\theta = \sqrt{n}(\mu - B)/\sigma =$ $\sqrt{n}IN(R)$, R being the fraction of the population to the right of B (i.e., the true reliability). Then

$$D(\gamma) = CN_\theta(\sqrt{n}IN(\beta) + IN(\gamma))$$

or

$$D(\gamma) = CN_o(\sqrt{n}IN(\beta) + IN(\gamma) - \sqrt{n}IN(R))$$

$$= CN_o(IN(\gamma) + \sqrt{n}(IN(\beta) - IN(R))).$$

The density of γ is found by differentiating $d(\gamma)$ with respect to γ, which yields

$$d(\gamma) = \frac{f(IN(\gamma) + \sqrt{n}(IN(\beta) - IN(R)))}{f(IN(\gamma))}$$

where

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Using this form of $f(x)$, the density of γ can be expressed as

$$d(\gamma) = \exp -\tfrac{1}{2}((IN(\gamma) + \sqrt{n}(IN(\beta) - IN(R)))^2 - (IN(\gamma\ ))^2)$$

or

$$d(\gamma) = \exp(c\sqrt{n}IN(\gamma) - nc^2/2) \quad , \quad 0 < \gamma < 1 \qquad (5)$$

where $$c = IN(R) - IN(\beta).$$

To check for the montonicity of $d(\gamma)$, the derivative of $d(\gamma)$ was examined to see when it is less than or greater than zero. The derivative is

$$d'(\gamma) = c\sqrt{n}d(\gamma)/f(IN(\gamma))$$

in which everything is positive, with the possible exception of the constant c. If $R > \beta$, then $IN(R) > IN(\beta)$ and so $c > 0$. In the case, $d'(\gamma) > 0$ and so $d(\gamma)$ is monotone increasing for all γ. In a similar fashion, it can be shown that if $R < \beta$ then $c < 0$ and $d'(\gamma) < 0$, in which case $d(\gamma)$ in monotone decreasing for all γ. Hence, if the true reliability R is greater than the desired reliability β, values of γ near 1.00 are the most likely; conversely, if R is less than β, values of γ close to 0.00 are most likely.

Since $IN(0) = -\infty$ and $IN(1) = \infty$, the value of $d(0)$ and $d(1)$ will be 0 or ∞, depending on the sign of c. Thus,

$$d(0) = \begin{cases} 0, & R > \beta \\ \infty, & R < \beta \end{cases}$$

$$d(1) = \begin{cases} \infty, & R > \beta \\ 0, & R < \beta \end{cases} .$$

The form of $d(\gamma)$ is reflected in the plots of $d(\gamma)$ in Figures 7 to 9.

A cursory examination of the plots of $d(\gamma)$ indicate that generally the same conclusions would hold as those for $d(\gamma)$ when σ is unknown. In the above example on tire reliability, it was shown that one can achieve a much higher confidence if σ is known. This is reflected in the composite plots in Figures 10 and 11 of the density of γ when σ is known and unknown. These results suggest that if σ is known, it is easier to draw a conclusion on whether R is greater than β since high values of γ are more likely than when σ is unknown.

The p^{th} moment of γ is given by

$$E(\gamma^p) = \int_0^1 \gamma^p d(\gamma) d\gamma$$

$$= \int_0^1 \frac{f(IN(\gamma) + \sqrt{n}c)\ d\gamma}{f(IN(\gamma))}$$

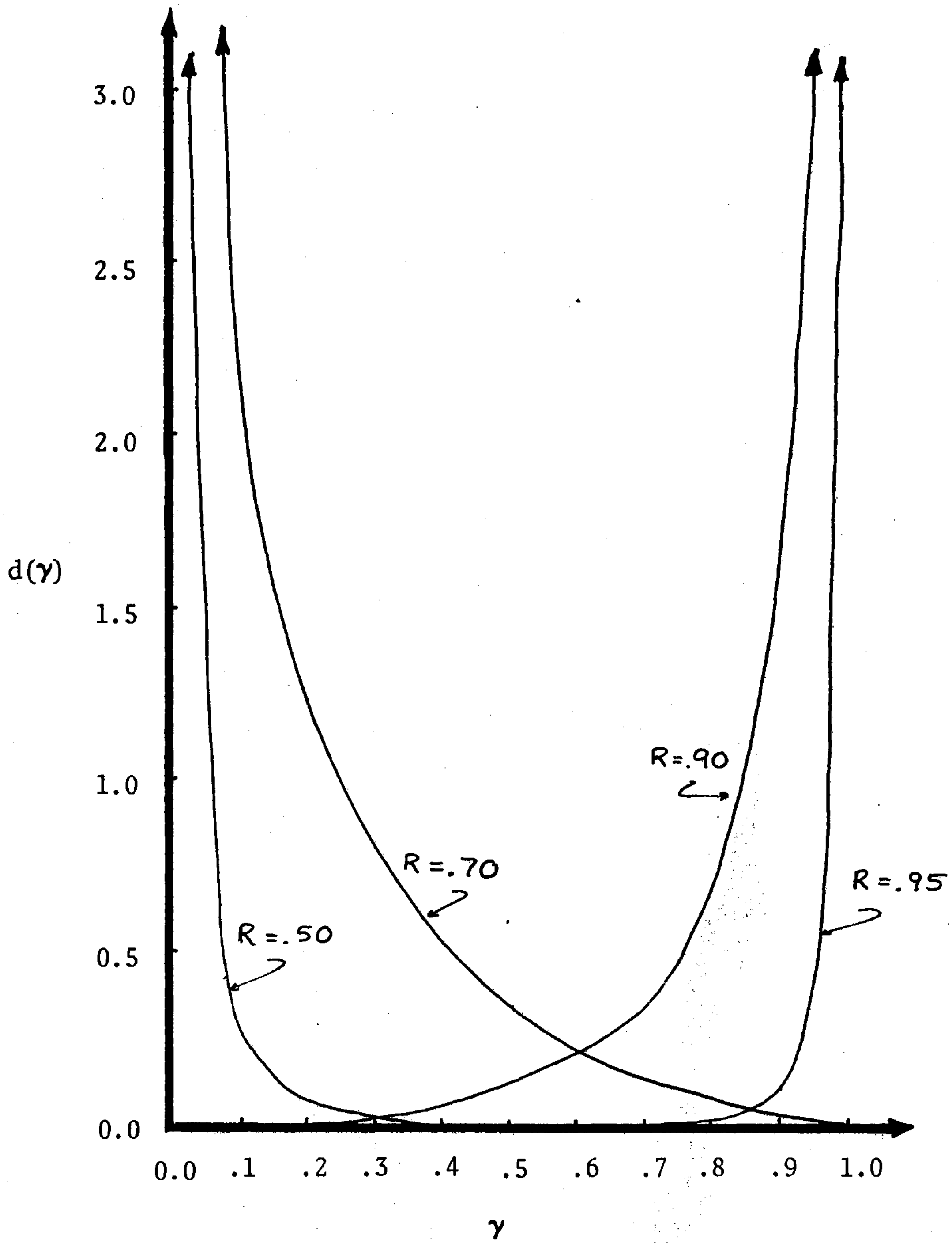


Figure 7. Densities of γ; Normal With σ Known; $\beta = .80$, $n = 20$

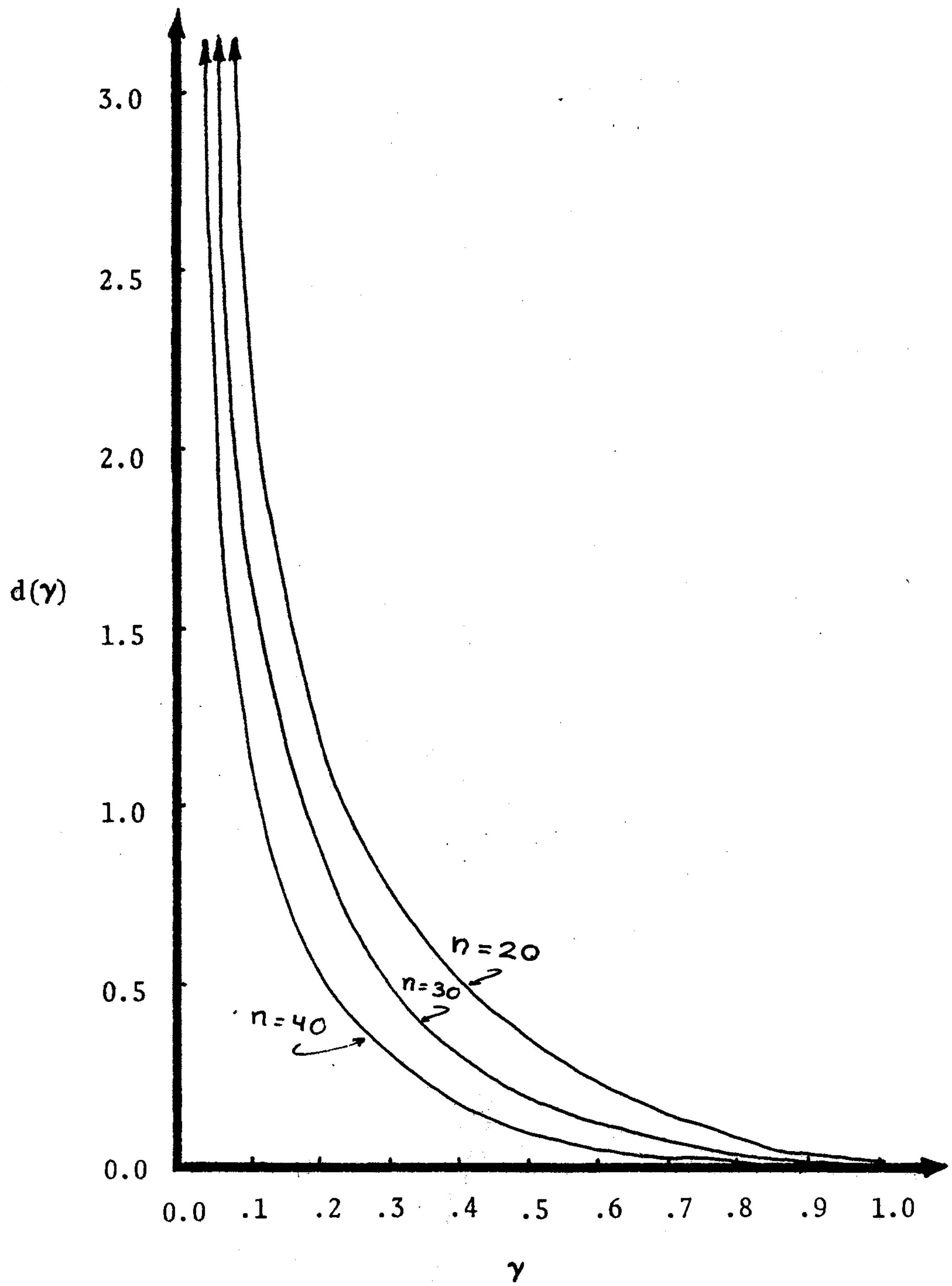


Figure 8. Densities of γ; Normal With σ Known; β = .80, R = .70

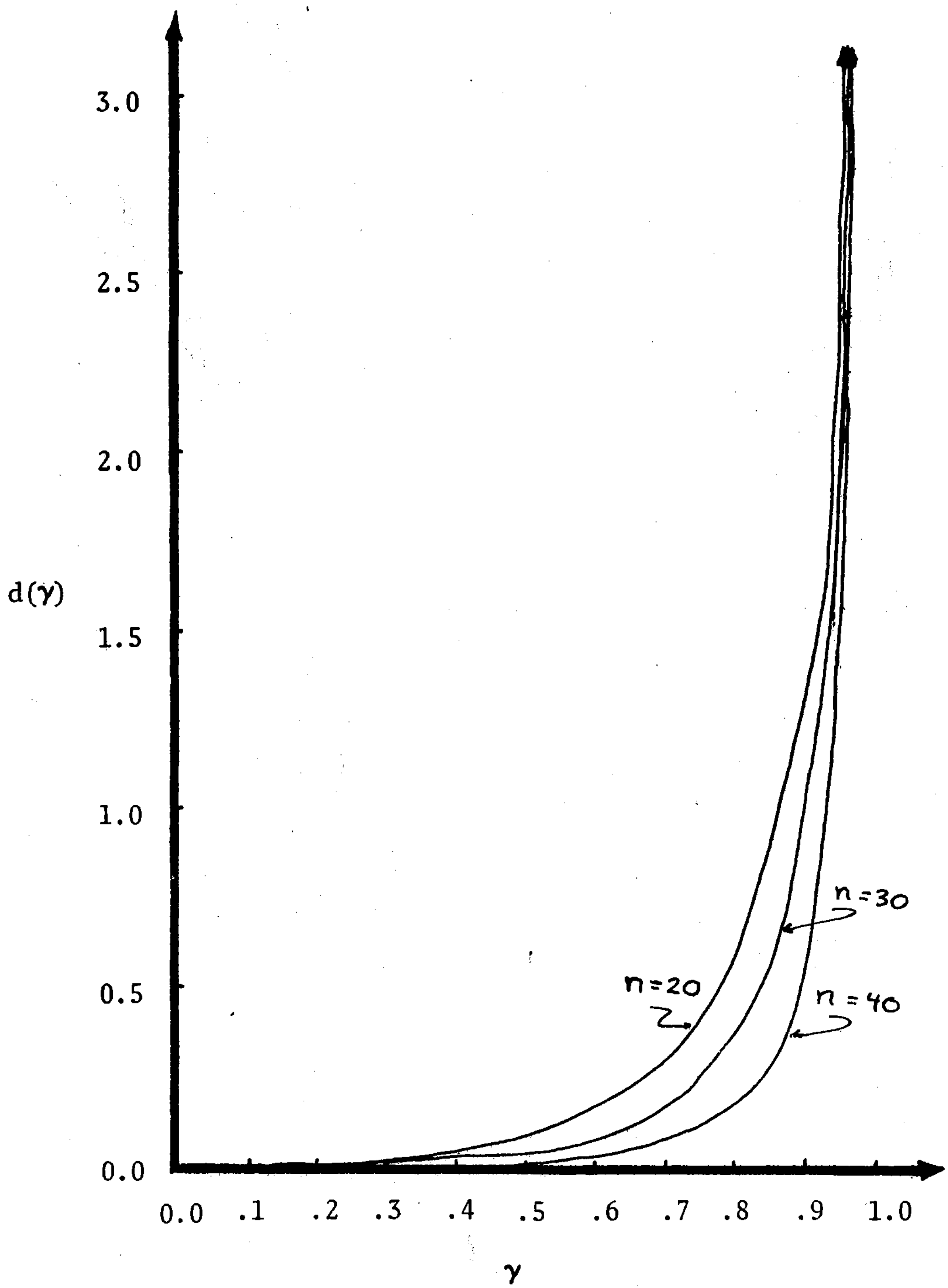


Figure 9. Densities of γ; Normal With σ Known;
$\beta = .80$, $R = .90$

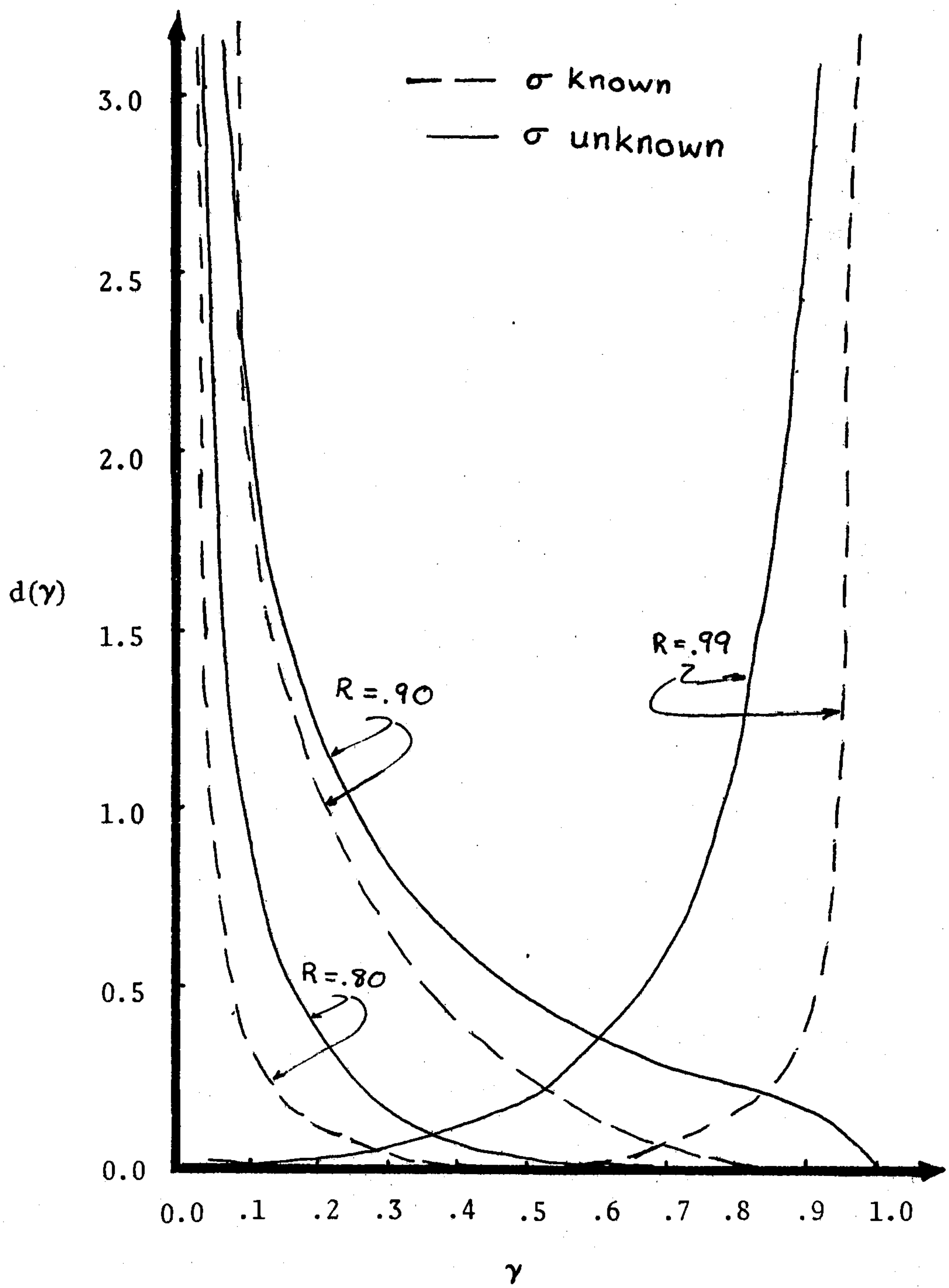


Figure 10. Densities of γ; Normal With σ Known and Unknown; β = .95, n = 20

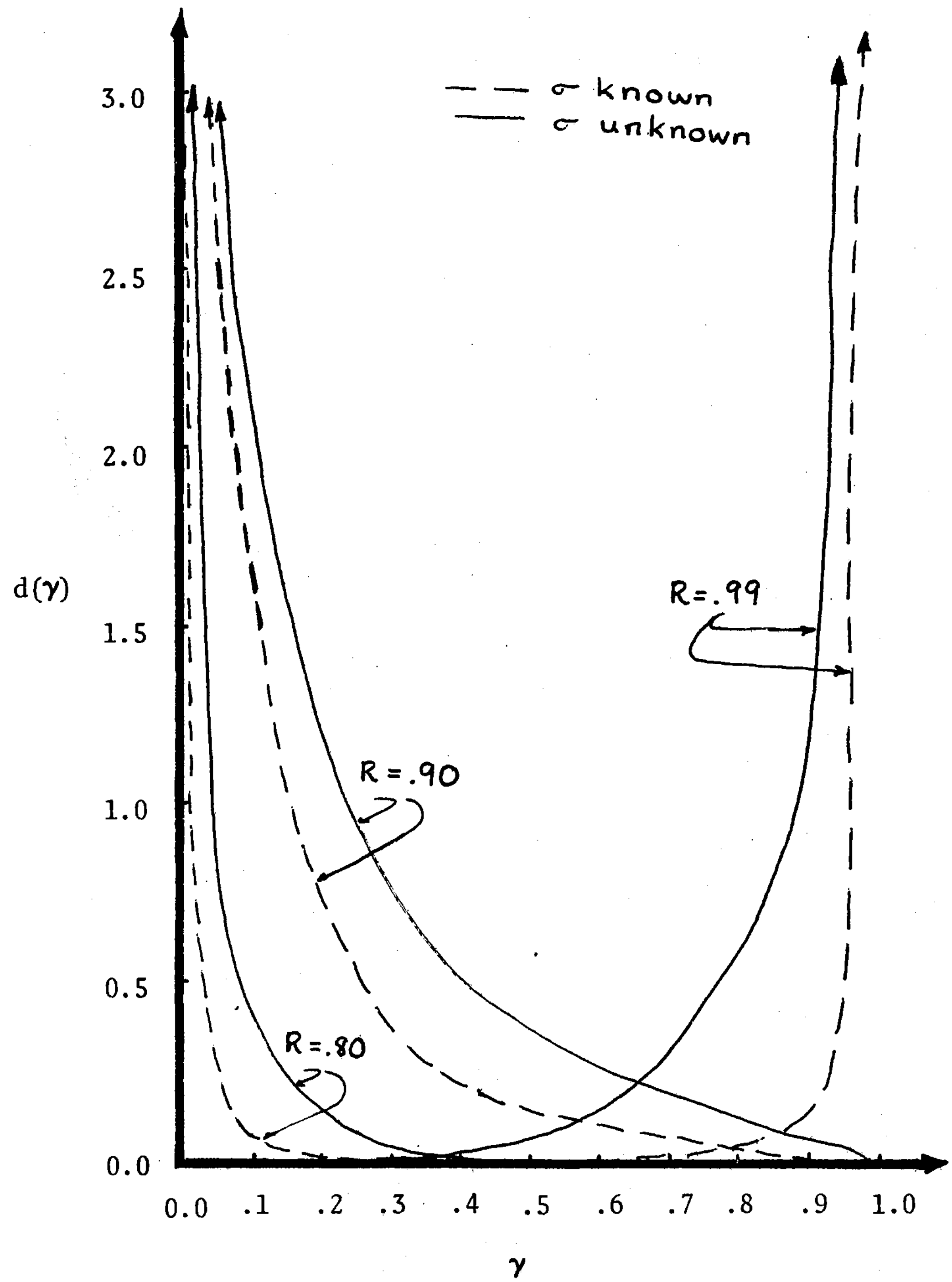


Figure 11. Densities of γ; Normal With σ Known and Unknown; $\beta = .95$, $n = 30$

where

$c = IN(R) - IN(\beta)$.

Let
$$x = IN(\gamma)$$
$$\text{or} \quad \gamma = CN_o(x)$$
$$d\gamma = f(x)dx.$$

Then

$$E(\gamma^p) = \int_{-\infty}^{\infty} (IN(x))^p f(x + \sqrt{n}c)dx.$$

This provides a non-closed form for $E(\gamma)$ and $Var(\gamma) = E(\gamma^2) - E^2(\gamma)$. No effort was made to tabulate values for this case. It is likely that such a compilation would show that $E(\gamma)$ would be closer to 1.00 and 0.00, as the case may be, with smaller variances in every case, than that seen when σ was unknown.

A Gamma Distribution

Let S be a random sample of size n from a population with a gamma distribution of the form

$$g_d(x) = x^{d-1}e^{-x}/\Gamma(d) \quad , \quad x > 0$$

where $d > 0$ is an unknown parameter. According to Bain and Weeks (13), one formula for a γ-probability lower OSTL on at least β of the population is

$$L(S) = G_q^{-1}(1 - \beta)$$

where

$G_q(\cdot)$ is the incomplete gamma with parameter q

q is the solution to

$$G_{nq}(\Sigma x) = \gamma. \tag{6}$$

If a QL is set equal to B, then the value of γ required to make $L(S) = B$ is given by (6), where q is the solution to

$$B = G_q^{-1}(1 - \beta)$$

or

$$G_q(B) = 1 - \beta.$$

This provides an implicit solution for the confidence γ.

The distribution function of γ is given by

$$D(\gamma) = \Pr(G_{nq}(\Sigma x) \leq \gamma) \quad , \quad 0 < \gamma < 1$$

$$= \Pr(\Sigma x \leq G_{nq}^{-1}(\gamma)).$$

But Σx is a gamma random variable with parameter nd. Thus,

$$D(\gamma) = G_{nd}(G_{nq}^{-1}(\gamma)) \quad , \quad 0 < \gamma < 1. \tag{7}$$

It would be desirable to replace d and B by a parameter R, the fraction of the population to the right of B. The first step is to consider the expression

$$\int_B^{\infty} g_d(x)dx = R$$

or

$$\int_0^B g_d(x)dx = G_d(B) = 1 - R.$$

This leaves the parameter d as an implicit function of R and B. It appears that $D(\gamma)$ will still have the QL of B as a parameter. This would make tabulation of results for $d(\gamma)$, $E(\gamma)$, and $Var(\gamma)$ less general since they would depend on the particular choice of B.

The density of γ is given by

$$d(\gamma) = \frac{g_{nd}(x)}{g_{nq}(x)} \qquad x = G_{nq}^{-1}(\gamma) \quad , \quad 0 < \gamma < 1.$$

Substituting the formula for the gamma density yields

$$d(\gamma) = (G_{nq}^{-1}(\gamma))^{n(d-q)} \ \frac{\Gamma(nq)}{\Gamma(nd)}. \tag{8}$$

Since $G_{nq}^{-1}(\gamma)$ is positive for all γ, then $d(\gamma)$ will be a monotone increasing (decreasing) function if $(d-q)$ is greater than (less than) zero. Suppose $R > \beta$. Then $1-R < 1-\beta$. Using the implicit expressions for d and q, it is apparent that

$$G_d(B) < G_q(B), \text{ for all } B$$

which implies

$$d > q.$$

Hence, if $R > \beta$, $d > q$ and so $d(\gamma)$ is monotone increasing. In a similar fashion, it can be shown that if $R < \beta$, $d(\gamma)$ is a monotone decreasing function of γ. This is the same type of result seen in the normal distribution problems.

Since $G_{nq}^{-1}(0) = 0$ and $G_{nq}^{-1}(1) = \infty$, the values of $d(0)$ and $d(1)$ will be zero or infinity, depending on the sign of $(d-q)$ in Equation (8). Considering the above findings on the sign of $(d-q)$ with respect to the order of R and β,

$$d(0) = \begin{cases} 0, & R > \beta \\ \infty, & R < \beta \end{cases}$$

$$d(1) = \begin{cases} \infty, & R > \beta \\ 0, & R < \beta \end{cases}.$$

Hence, $d(\gamma)$ begins at zero and goes to infinity for all choices of R and β. The only matter is whether $d(\gamma) = 0$ for $\gamma = 0$ or $\gamma = 1$. This is determined by the order of R and β.

The p^{th} moment of γ is given by

$$E(\gamma^p) = \int_0^1 \gamma^p d(\gamma) d\gamma$$

$$\int_0^1 \gamma^p \frac{g_{nd}\left[G_{nq}^{-1}(\gamma)\right]}{g_{nq}\left[G_{nq}^{-1}(\gamma)\right]} d\gamma.$$

To perform this integration, let

$$x = G_{nq}^{-1}(\gamma)$$

or

$$\gamma = G_{nq}(x)$$

$$d\gamma = G_{nq}(x)\ dx.$$

Then

$$E(\gamma^p) = \int_0^\infty G_{nq}^p(x) g_{nd}(x) dx.$$

This provides a non-closed form for $E(\gamma)$ and $Var(\gamma) = E(\gamma^2) - E^2(\gamma)$. As stated above, tabulation of results would be quite ungeneral since the value of d would depend on the choice of B. Removal of B as a parameter would be a useful area of further research on this problem. The other alternative would be to find another formula for the tolerance limit which would have "good" statistical properties, and be easy to use.

A Weibull Distribution

Let S be a random sample of size n from a population with a Weibull distribution of the form

$$w_a(x) = (b/a^b)x^{b-1}\exp-(x/a)b \quad , \quad x > 0$$

in which $b > 0$ is known and $a > 0$ is unknown. According to Bain and Weeks (13), one formula for a γ-probability lower OSTL on at least β of the population is

$$L(S) = (-\Sigma x^b \ell n\, \beta / G_n^{-1}(\gamma))^{1/b}$$

where

$G_n(\cdot)$ = incomplete gamma with parameter n.

Suppose a QL = B is set by the manufacturer. The required value of γ to make L(S) = B (i.e., the observed confidence) is found by solving

$$B = (-\Sigma x^b \ell n \beta / G_n^{-1}(\gamma))^{1/b}$$

which yields

$$\gamma = G_n(-\Sigma x^b \ell n \beta / B^b).$$

This provides a formula for the observed confidence γ for any sample of size n, for desired reliability β.

As an example of how this might be used, suppose a manufacturer of high intensity lamps wishes to know with what confidence he can claim that at least 95% of his lamps will last over 15 days (i.e., have 95% reliability). To do this, he first put ten lamps on test with the following results:

Lamp #	1	2	3	4	5	6	7	8	9	10	
lifetime	18	17	24	18	26	25	23	23	19	20	(days)

The manufacturer knows from experience that light bulb lifetimes are Weibull-distributed and that b = 2. Using this data, the expression for γ becomes

$$\gamma = G_{10}(2.112) \doteq .01.$$

The value of γ is found by use of any tables on the incomplete gamma with parameter n = 10. Hence, the manufacturer can claim his lamps are 95% reliable with about 1% confidence. The lower OSTL with 95% confidence is 3.89.

The distribution function of γ is given by

$$D(\gamma) = \Pr(G_n(-\Sigma x^b \ln\beta/B^b) \leq \gamma) \quad , \quad 0 < \gamma < 1$$

$$= \Pr(\Sigma x^b \leq - B^b G_n^{-1}(\gamma)/\ln\beta)$$

$$= \Pr(\Sigma(x/a)^b \leq - (B/a)^b G_n^{-1}(\gamma)/\ln\beta)$$

by dividing through by a^b. But $\Sigma(x/a)^b$ is a gamma random variable with parameter n. Therefore

$$D(\gamma) = G_n(-(B/a)^b G_n^{-1}(\gamma)/\ln\beta). \tag{9}$$

Note that if a fraction R of the Weibull population lies above B, this is expressed by

$$\int_B^\infty ba^{-b}x^{b-1}\exp-(x/a)^b dx = R$$

or

$$\exp-(B/a)^b = R$$

or

$$-(B/a)^b = \ell n R.$$

Substituting this into (9) yields

$$D(\gamma) = G_n((\ell n R/\ell n \beta)G_n^{-1}(\gamma)), \quad 0 < \gamma < 1.$$

The density of γ is given by

$$d(\gamma) = \frac{\ell n R}{\ell n \beta} \frac{g_n((\ell n R/\ell n \beta)G_n^{-1}(\gamma))}{g_n(G_n^{-1}(\gamma))}.$$

Substituting the formula of the gamma density into this expression yields

$$d(\gamma) = \frac{\ell n R}{\ell n \beta} \exp((1 - \ell n R/\ell n \beta)G_n^{-1}(\gamma)).$$

Since $d(\gamma)$ is essentially e raised to a power, $d(\gamma)$ is monotone increasing or decreasing depending only on the sign of the exponent. Suppose $R > \beta$. Then $\ell n R > \ell n \beta$, and $\ell n R/\ell n \beta < 1$, since the log of R and β will be negative. Then

$$1 - \ell n R/\ell n \beta > 0$$

and so $d(\gamma)$ will be monotone increasing. In a similar fashion, it can be shown that if $R < \beta$, then $d(\gamma)$ is monotone decreasing.

The values of $d(\gamma)$ for $\gamma = 0$ and $\gamma = 1$ are found by first noting that $G_n^{-1}(0) = 0$ and $G_n^{-1}(1) = \infty$. The values of $d(0)$ and $d(1)$ will be determined by the sign of the exponent in the second term of $d(\gamma)$. Thus,

$$d(0) = (\ln R/\ln \beta)^n \text{ for all } R \text{ and}$$

$$d(1) = \begin{cases} \infty, & R > \beta \\ 0, & R < \beta \end{cases}.$$

These limits are partially illustrated in the plots of $d(\gamma)$ in Figures 12 through 15. The values of $G_n^{-1}(\gamma)$ were found by using the approximation

$$G_n^{-1}(\gamma) \doteq IN(\gamma) + \sqrt{4n-1}.$$

These graphs show a very marked skewness in $d(\gamma)$, perhaps even more so than that seen in either of the normal problems considered earlier. This is at least partially verified in the composite plots of $d(\gamma)$ in Figure 16. In these figures, the graphs of $d(\gamma)$ are given for the same sample size n, R, and β for the Weibull, normal with σ known and with σ unknown. It would appear that γ would be a more sensitive test for R being greater than β if the data is known to be Weibully distributed and γ is computed using the Weibull formula. As an illustration, suppose one were to appeal to the Central Limit Theorem and compute γ using Equation (1) (i.e., for the normal with σ unknown). Using the data in the lamp life example above, the observed confidence is found to be .45, instead of the value .01 when using the Weibull formula. This would

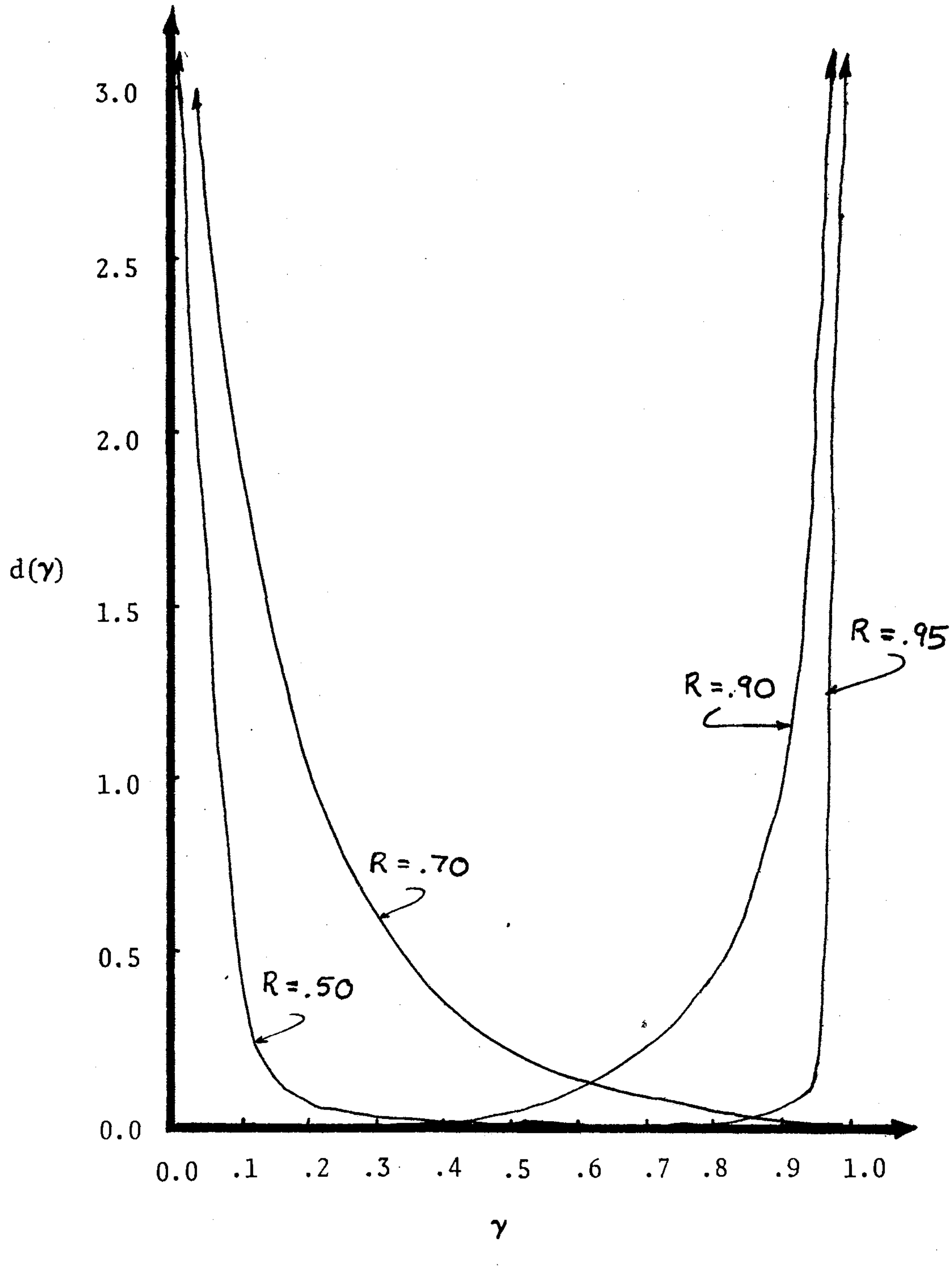


Figure 12. Densities of γ; Weibull; β = .80, n = 10

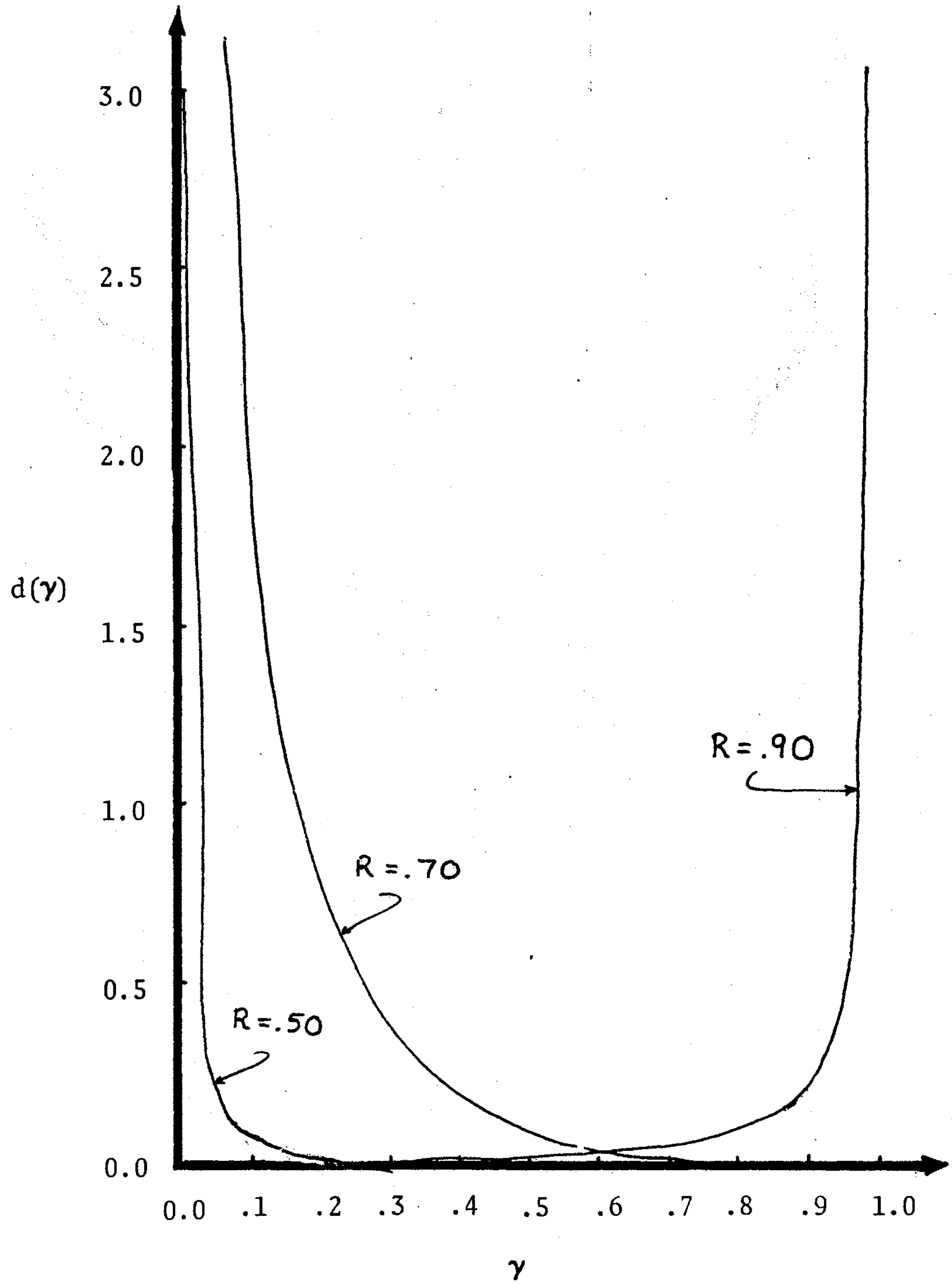


Figure 13. Densities of γ; Weibull; $\beta = .80$, $n = 20$

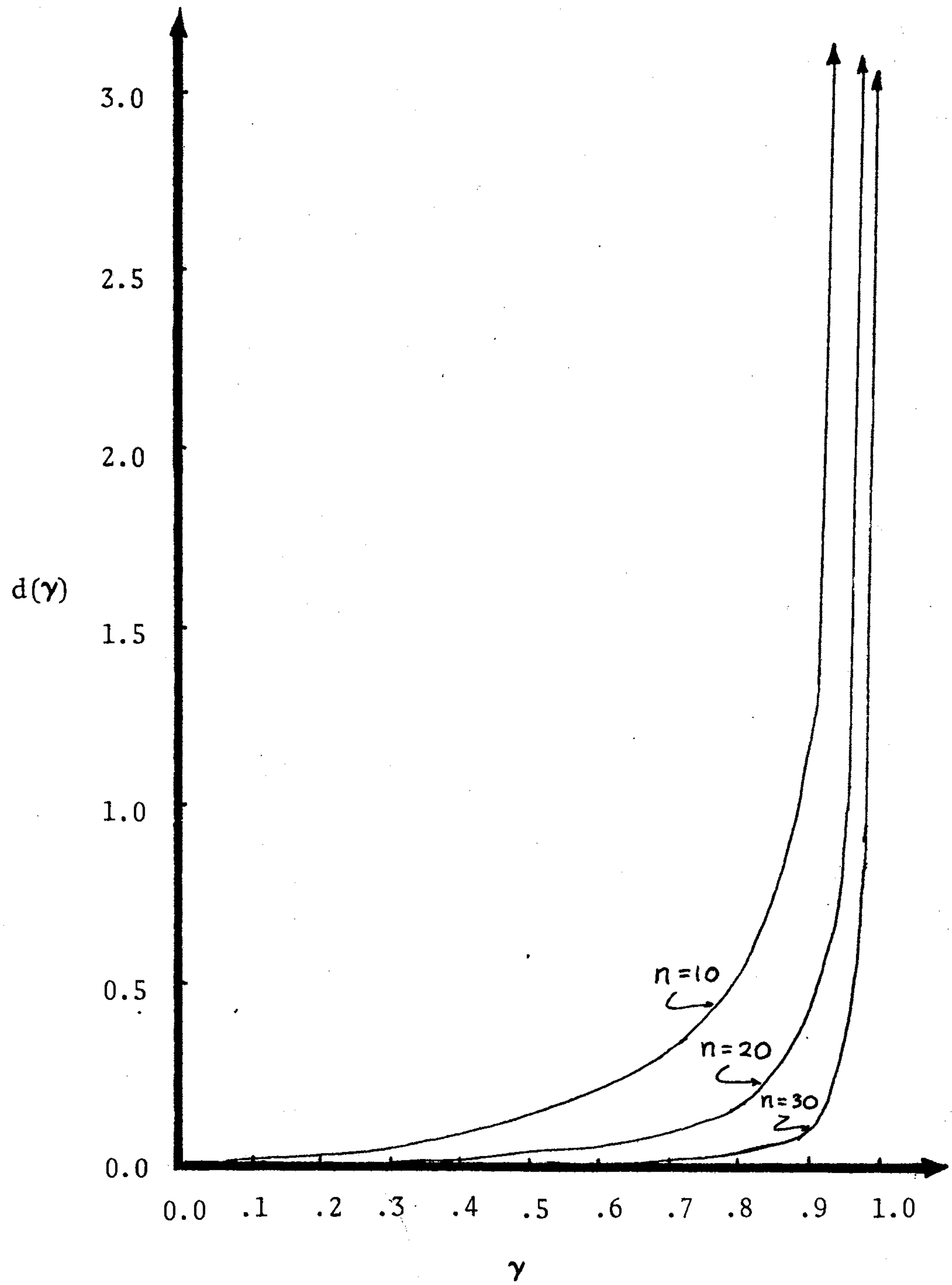


Figure 14. Densities of γ; Weibull; β = .50, R = .70

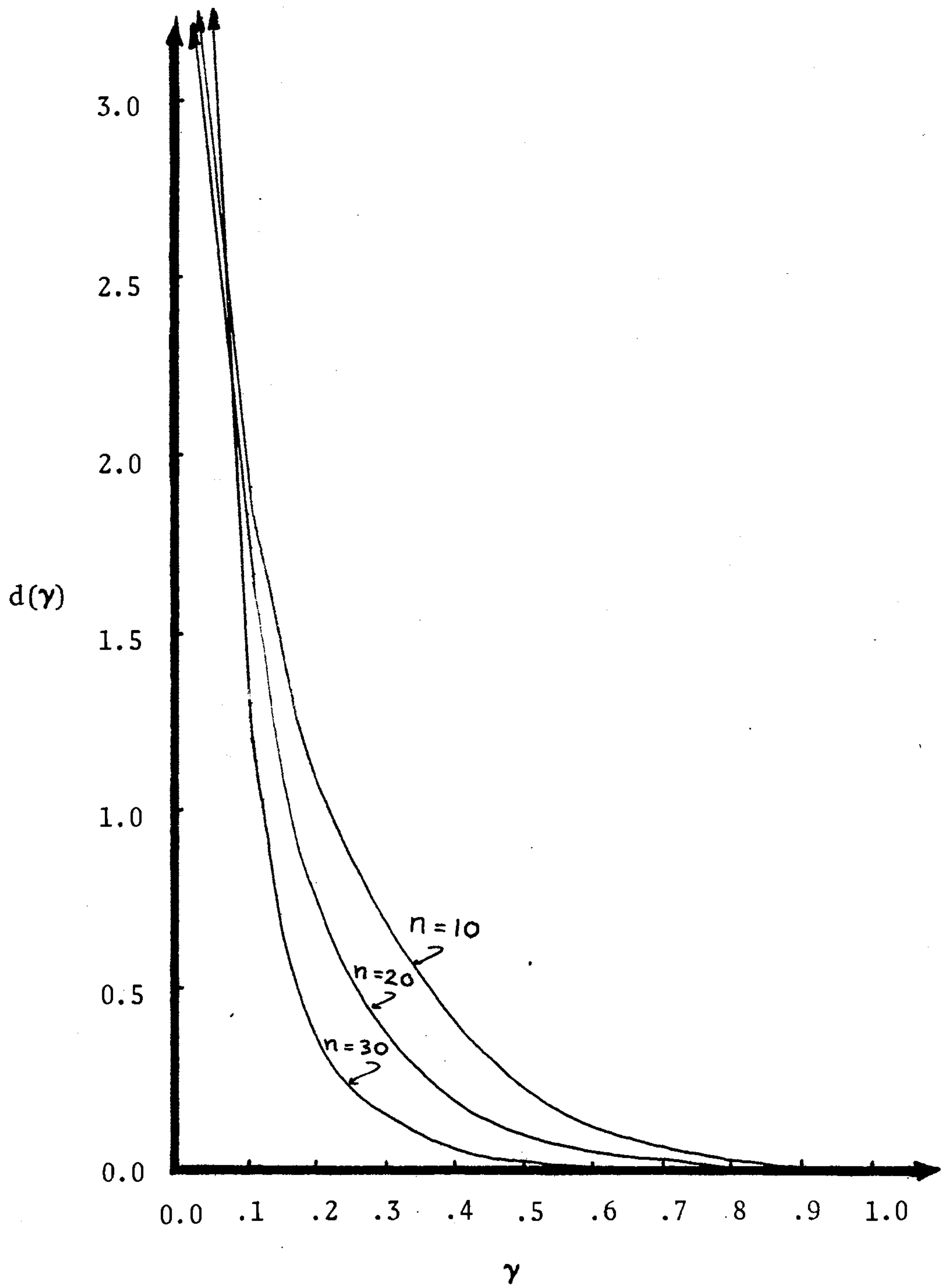


Figure 15. Densities of γ; Weibull; $\beta = .80$, $R = .70$

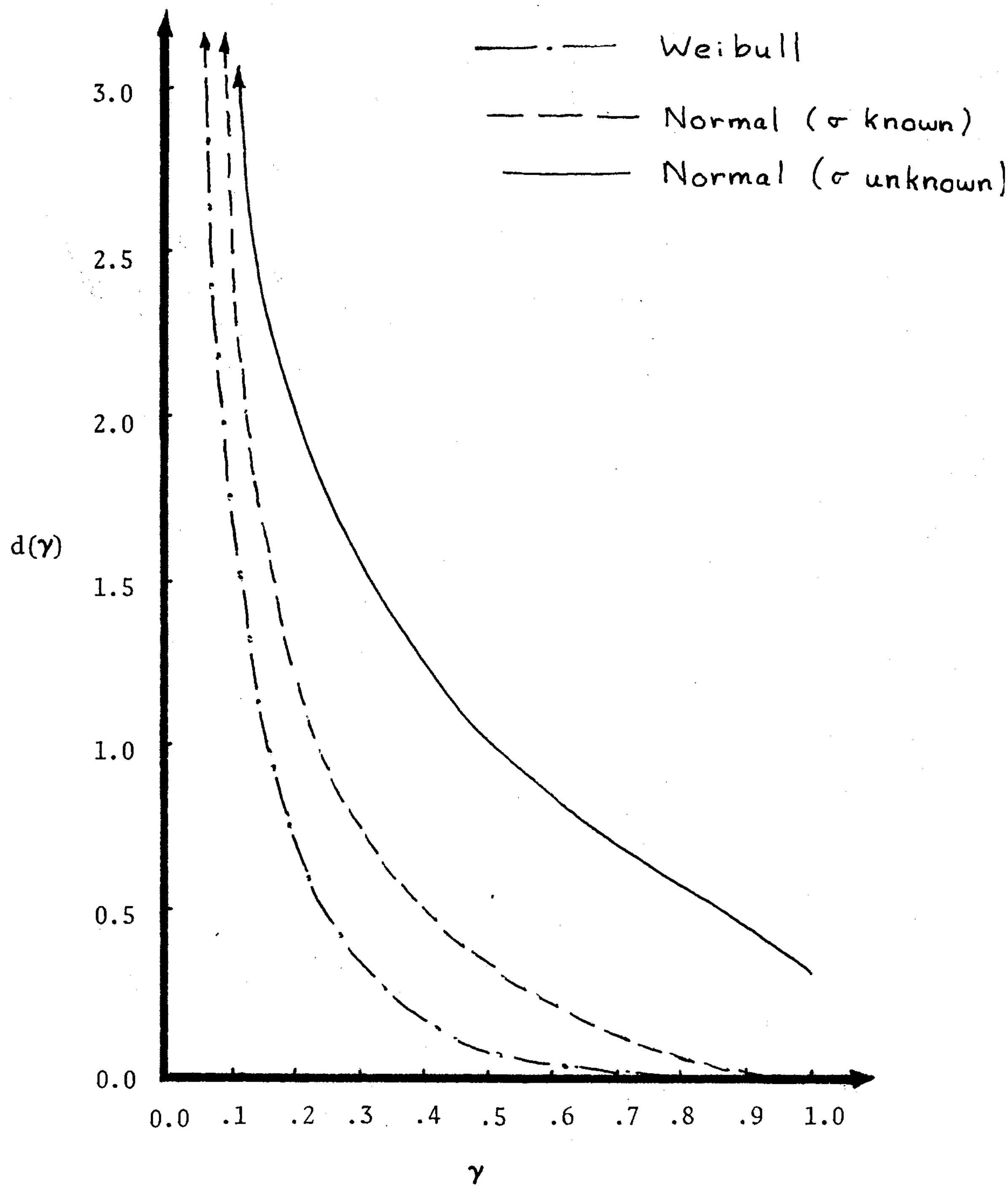


Figure 16. Densities of γ; Weibull and Normal With σ Known and Unknown; $\beta = .80$, $R = .70$, $n = 10$

suggest that it would be quite worthwhile to determine the true model for the data, rather than appeal to the Central Limit Theorem as casually as many texts seem to suggest.

To complete the presentation on the Weibull, the p^{th} moment is developed here. The p^{th} moment is given by

$$E(\gamma^p) = \int_0^1 \gamma^p d(\gamma) d\gamma$$

or

$$E(\gamma^p) = \int_0^1 \gamma^p r \frac{g_n(rG_n^{-1}(\gamma))}{g_n(G_n^{-1}(\gamma))} d\gamma$$

where $r = \ell n R / \ell n \beta$. To integrate this more easily, let

$$x = G_n^{-1}(\gamma)$$

or

$$\gamma = G_n(x)$$

$$d\gamma = g_n(x) dx.$$

Then

$$E(\gamma^p) = r\int_0^\infty G_n^p(x) g_n(rx) dx$$

or

$$E(\gamma^p) = \frac{\ell n R}{\ell n \beta} \int_0^\infty G_n^p(x) g_n(x \ell n R / \ell n \beta) dx.$$

This provides a non-closed form for obtaining $E(\gamma)$ and $Var(\gamma)$. This writer feels that this would be an easier form to work with than the form involving the direct integration of γ since finding inverse points of an incomplete gamma may be more difficult to find than values of the incomplete gamma itself.

An Exponential Distribution

Let S be a random sample of size n from a population with an exponential distribution of the form

$$h_a(x) = \frac{1}{a} e^{-x/a} \quad , \quad x > 0$$

where $a > 0$ is unknown. This density is a special case of the Weibull density in which $b = 1$. Using the results from the Weibull yields

$$\gamma = G_n(-\Sigma x \ln \beta / B)$$

as the computational form for the observed confidence γ.

The distribution function of γ for the Weibull population is

$$D(\gamma) = G_n\left(\frac{\ln R}{\ln \beta} G_n^{-1}(\gamma)\right)$$

which does not depend on the known parameter b in the Weibull. While this result might at first seem surprising, it is no different than the apparent loss of σ as a parameter in the normal case where σ was known. The parameter σ was simply incorporated into the parameter R, along with the QL of B. As a result, one can use the distribution function, density, and moments of γ for the Weibull as those for the exponential.

CHAPTER IV

SOME RESULTS IN ATTRIBUTE TESTING

The results of the prior chapters have utilized the measurement of a variable, such as time to failure. That approach is referred to as a variables testing procedure. Such testing requires a knowledge of the form of the variable's probability distribution.

However, this distribution form is often unknown, particularly in exploratory testing. To avoid making any incorrect distributional assumptions, one can work nonparametrically with the number of "failures" in the n samples on test. This approach is referred to as attribute testing and will be the type of testing considered in this chapter.

A Binomial Model Result

Let S be a random sample of size n from an unspecified population. Let N(S) be the number of sample values greater than a specified QL, say B. That is, N(S) is the number of items on test which do not "fail". It is desired to express as a function of N(S) the confidence γ that at least β of the population lies above B.

Let R be the fraction of the population above B. Then $\underline{R}$ is a γ-confidence lower limit on R (see (16)) if

$$\gamma = \sum_{i=N(S)}^{n} \binom{n}{i} \underline{R}^{i}(1-\underline{R})^{n-1}. \qquad (1)$$

The question of interest is: what must γ be so that $\underline{R} = \beta$ is a γ-confidence lower limit on R? This value of γ is obtained by letting $\underline{R} = \beta$ and computing γ from Equation (1).

$$\gamma = \sum_{i=N(S)}^{n} \binom{n}{i} \beta^{i}(1 - \beta)^{n-1}. \tag{2}$$

The left side of (2) can be evaluated by the incomplete beta function

$$Ib(N(S);\beta = \frac{\Gamma(n+1)}{\Gamma(N(S))\Gamma(n - N(S) + 1)} \int_0^{\beta} u^{N(S)-1}(1+u)^{n-N(S)}du.$$

Hence,

$$\gamma = Ib(N(S);\beta).$$

This provides a computational form for the observed confidence coefficient γ for a given N(S) and β.

The distribution function of γ is

$$D(\gamma) = Pr(Ib(N(S);\beta) \leq \gamma) \quad , \quad 0 < \gamma < 1$$

$$= Pr(N(S) \leq [Ib^{-1}(\gamma;\beta)])$$

where

[x] is the largest integer $\leq$ x.

But N(S) is a binomial random variable with parameters n and R. Thus,

$$D(\gamma) = \sum_{i=0}^{[Ib^{-1}(\gamma;\beta)]} \binom{n}{i} R^{i}(1-R)^{n-1}$$

$$= 1 - Ib([Ib^{-1}(\gamma;\beta)] + 1;\ R), \quad 0 < \gamma < 1.$$

The probability mass function of γ is

$$d(\gamma) = \binom{n}{x} R^x (1 - R)^{n-x} \Big|_{x = Ib^{-1}(\gamma;\beta)} \quad , \quad 0 < \gamma < 1.$$

This is easily obtained by writing the incomplete beta in its equivalent form as a cumulative binomial. Preliminary work with the probability mass function indicate that as R moves away from β, the mode of the function moves towards either $\gamma = 0.0$ or 1.0, depending in the usual manner on whether R is greater than or less than β. While no work has been done on the effect of sample size, it would not be surprising if the movement of the mode towards the end points would be accentuated by an increasing sample size, as well as an increase in probability of the mode.

A Poisson Model Result

Suppose n randomly chosen items are placed on test for a time T and replaced as they fail. This is often done in tests of electronic systems made up of n components so that the system remains operational. If the failure rate of the items is a constant value λ, then the failures are said to follow a Poisson process (16). The number of failures in time T, say N(S), is a Poisson random variable with parameter $n\lambda T$.

The reliability R of the products on test is given in (16) as

$$R = \exp(-\lambda T).$$

It is desired to find a γ-confidence lower limit on R, say $\underline{R}(S)$. The value of γ for which $\underline{R}(S) = \beta$ would give the confidence with which one could say the product is 100β% reliable.

For a sample S of size n, let

$$\gamma = \Pr(R > \underline{R}(S)).$$

This is the same as saying

$$\gamma = \Pr(e^{-\lambda T} > \underline{R}(S))$$

$$= \Pr(-\lambda T > \ell n \underline{R}(S))$$

$$= \Pr(\lambda T < -\ell n \underline{R}(S)).$$

The function $-\ell n R(S)$ is now a γ-confidence upper limit on λT for a given value of S. Then

$$\gamma = \sum_{x=0}^{N(S)-1} \frac{(-n\ell n\ \underline{R}(S))^x e^{-n\ell n\ \underline{R}(S)}}{x!}. \tag{3}$$

But in reference (16), it is shown that Equation (3) can be expressed as

$$\gamma = 1 - G_{N(S)-1}(-n\ell n\ \underline{R}(S)) \tag{4}$$

where $G_a(\cdot)$ is the incomplete gamma with parameter a. Solving Equation (4) yields

$$-\ell n\ \underline{R}(S) = G^{-1}_{N(S)-1}(1-\gamma)/n.$$

Then the value of γ such that $\underline{R}(S) = \beta$) is

$$\gamma = 1 - G_{N(S)-1}(-n\ell n\ \beta)$$

or

$$\gamma = CP_{-n\ell n\beta}(N(S))$$

where $CP_a(\cdot)$ is the cumulative Poisson with parameter a. This provides a computational form for the observed confidence for a given desired reliability β and number of failures equal to N(S).

The distribution function of γ is

$$D(\gamma) = \Pr(CP_{-n\ln\beta}(N(S)) \leq \gamma), \quad 0 < \gamma < 1$$

$$= \Pr\left(N(S) \leq \left[CP^{-1}_{-n\ln\beta}(\gamma)\right]\right).$$

But N(S) is a Poisson random variable with parameter nλT. Thus,

$$D(\gamma) = CP_{n\lambda T}\left(\left[CP^{-1}_{-n\ln\beta}(\gamma)\right]\right), \quad 0 < \gamma < 1.$$

But if

$$R = e^{-\lambda T}$$

then

$$\lambda T = -\ln R.$$

Thus,

$$D(\gamma) = CP_{-n\ln R}\left(\left[CP^{-1}_{-n\ln\beta}(\gamma)\right]\right), \quad 0 < \gamma < 1.$$

The distribution function of γ is now expressed in terms of the sample size n, desired reliability β, and actual reliability R.

The probability mass function is given by

$$d(\gamma) = \left.\frac{e^{-n\ln R}(-n\ln R)^x}{x!}\right|_{x = CP^{-1}_{-n\ln\beta}(\gamma)}, \quad 0 < \gamma < 1.$$

Little work has been done on plotting d(γ), but one would anticipate that the mode of d(γ) would shift drastically towards 0.00 or 1.00, depending in the usual manner on whether R was greater or less than β. The effect of increasing sample size should be to accentuate the shift and increase the probability of the mode.

CHAPTER V

A REVISED OSTL FOR THE NORMAL DISTRIBUTION

In Chapter III, the OSTL for a normal population with σ unknown was

$$L(S) = \bar{x} - T_{\delta}^{-1}(\gamma)s/\sqrt{n}. \qquad (1)$$

The question arises as to whether L(S) should ever be greater than $\bar{x}$. For example, it is possible that $\bar{x}$ could equal 17.3 while L(S) would equal 26.8. It would seem reasonable to have $\bar{x}$ serve as the lower OSTL in such an event since one would seem to be sure of covering at least β of the population by using the minimum of the values of $\bar{x}$ and L(S). To do this, the form of L(S) suggests that $T_{\delta}^{-1}(\gamma)$ should always be greater than or equal to zero, i.e.,

$$T_{\delta}^{-1}(\gamma) \geq 0 \quad \text{for all } \gamma.$$

But if this is true, then

$$\gamma \geq T_{\delta}(0).$$

What effect this type of restriction would have on the results of Chapter III is the subject of this chapter.

If L(S) = B, a specified QL, solving (1) for $T_{\delta}^{-1}(\gamma)$ yields

$$T_{\delta}^{-1}(\gamma) = \sqrt{n}(\bar{x} - B)/s.$$

If

$$T_\delta^{-1}(\gamma) \geq 0$$

then

$$\bar{x} > B.$$

Should $\bar{x}$ be less than B, it would seem reasonable to set

$$T_\delta^{-1}(\gamma) = 0$$

or

$$\gamma = T_\delta(0).$$

Then γ takes on the values of

$$\gamma = \begin{cases} T_\delta(\sqrt{n}(\bar{x} - B)/s) \ , & \bar{x} \geq B \\ T_\delta(0) \ , & \bar{x} < B \end{cases} .$$

The distribution function of γ will be found by applying the identity

$$D(\gamma) \equiv D(\gamma|\bar{x} \geq B)\Pr(\bar{x} \geq B) + D(\gamma|\bar{x} < B)\Pr(\bar{x} < B).$$

Now

$$D(\gamma|\bar{x} \geq B) = \Pr(T_\delta(\sqrt{n}(\bar{x} - B)/s \leq \gamma|\bar{x} \geq B)$$

$$= \Pr(\sqrt{n}(\bar{x} - B)/s \leq T_\delta^{-1}(\gamma)|\bar{x} \geq B).$$

But $\bar{x} \geq B$

is also expressible as

$$\sqrt{n}(\bar{x} - B)/s \geq 0.$$

Thus

$$D(\gamma|\bar{x} \geq B) = \frac{\Pr(\sqrt{n}(\bar{x} - B)/s \leq T_\delta^{-1}(\gamma),\ \sqrt{n}(\bar{x} - B)/s \geq 0)}{\Pr(\sqrt{n}(\bar{x} - B)/s \geq 0)}.$$

Since $\sqrt{n}(\bar{x} - B)/s$ is a noncentral-t random variable with parameter $\theta = \sqrt{n}\,\mathrm{IN}(R)$ and $n-1$ d.f.,

$$D(\gamma|\bar{x} \geq B) = \begin{cases} 0 & ,\ T_\delta^{-1}(\gamma) \leq 0 \text{ or } \gamma \leq T_\delta(0) \\ (T_\theta(T_\delta^{-1}(\gamma)) - T_\theta(0))/(1 - T_\theta(0)), & \gamma > T_\delta(0) \end{cases}.$$

In addition,

$$\begin{aligned} D(\gamma|\bar{x} < B) &= \Pr(T_\delta(0) \leq \gamma|\bar{x} < B) \\ &= \Pr(T_\delta(0) \leq \gamma) \\ &= \begin{cases} 0, & \gamma < T_\delta(0) \\ 1, & \gamma \geq T_\delta(0) \end{cases}. \end{aligned}$$

Using the fact that

$$\begin{aligned} \Pr(\bar{x} < B) &= \Pr(\sqrt{n}(\bar{x} - B)/s < 0) \\ &= T_\theta(0). \end{aligned}$$

$D(\gamma)$ can be expressed as

$$D(\gamma) = \begin{cases} 0 & ,\ \gamma < T_\delta(0) \\ T_\theta(0) & ,\ \gamma = T_\delta(0). \\ T_\theta(T_\delta^{-1}(\gamma)), & \gamma > T_\delta(0) \end{cases}$$

The distribution of γ is a mixed distribution with a mass point at

$T_\delta(0)$ and a density for $\gamma > T_\delta(0)$. Thus, the probability function of γ is

$$d(\gamma) = \begin{cases} 0 & , \gamma < T(0) \\ \dfrac{t_\theta(T_\delta^{-1}(\gamma))}{t_\delta(T_\delta^{-1}(\gamma))} & , \gamma > T(0) \end{cases}$$

and

$$\Pr(\gamma = T_\delta(0)) = T_\theta(0).$$

Plots of $d(\gamma)$ are the same as those for the unrestricted OSTL except that $d(\gamma)$ equals zero when $\gamma < T_\delta(0)$ and has a mass point at $T_\delta(0)$.

This approach to revising the OSTL does not increase the frequency of values of γ above $T_\delta(0)$; it only excludes low values of γ. What practical value this type of OSTL might have for the manufacturer is not immediately apparent since reporting a value of $\gamma = T_\delta(0)$ tells little except that $\bar{x}$ was less than B. If this should be unexpected under the circumstances, then perhaps this result is slightly useful.

CHAPTER VI

TWO-SIDED TOLERANCE LIMITS

In the previous chapters, one formula was used to compute a γ- probability OSTL on β of the population. Another common practice in industry is to use two formulas to compute limits between whcih β of the population will lie with probability γ. This set of formulas are called two-sided tolerance limits. The tolerance limits on the normal distribution derived by Wald and Wolfowitz (7) appear to be the ones most in use today.

Wald-Wolfowitz Limits

Let S be a random sample of size n from a normally distributed population with unknown mean and variance. Then, according to Wald and Wolfowitz,

$$U(S) = \bar{x} + ks$$

and

$$L(S) = \bar{x} - ks$$

are γ-probability tolerance limits on the center β of the population if

$$k = r\sqrt{\frac{f}{IC\ (1 - \beta)}}$$

where

$f = n - 1$

IC(p) = point on the chi-square with f d.f. below which a proportion p lies.

$$\gamma = \int_{1/\sqrt{n}-r}^{1/\sqrt{n}+r} dN(0,1).$$

To simplify the following presentation, the following notation is introduced. Let

$$\gamma = \int_{1/\sqrt{n}-r}^{1/\sqrt{n}+r} dN(0,1) = \Delta(r)$$

then

$$r = \Delta^{-1}(\gamma)$$

and

$$k = \Delta^{-1}(\gamma)\sqrt{\frac{f}{IC\ (1-\beta)}}.$$

It should be noted that U(S) and L(S) form only approximate tolerance intervals; however, the error is insignificant for large n (20 or more).

Suppose that upper and lower QL's of U and L, respectively, are assigned. Setting U(S) = U and L(S) = L, one finds two values of γ, denoted by γ_L and γ_U, to be

$$\gamma_U = \Delta\left\{\sqrt{\frac{IC\ (1-\beta)}{n}}\left(\frac{U-\bar{x}}{s}\right)\right\} \tag{1}$$

and

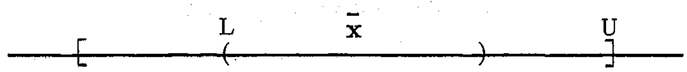
$$\gamma_L = \Delta\left\{\sqrt{\frac{IC\ (1-\beta)}{n}}\left(\frac{\bar{x}-L}{s}\right)\right\}. \tag{2}$$

Unless U and L are symmetric about $\bar{x}$, that is such that

$$\bar{x} = \frac{(U + L)}{2}$$

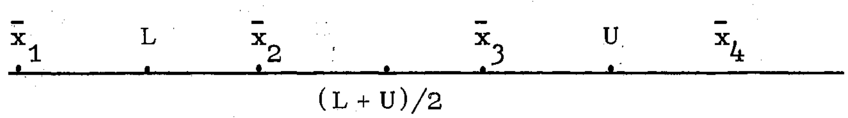
the values γ_U and γ_L will not be equal.

Consider the following example of such a situation.



If γ_L is used as the observed confidence, then the tolerance limits for such a choice of γ are shown by the parentheses. If γ_U is used as the observed confidence, the tolerance limits in that case are shown by the brackets. It appears that γ_U is the minimum confidence which allows for coverage of both QL's. Clearly, this is the larger of the confidences computed in (1) and (2). By inspecting the above drawing, it is also apparent that the larger confidence is the one associated with $\max(|\bar{x} - L| , |U - \bar{x}|)$. Hence, if $|U - \bar{x}|$ is largest, γ is given by (1); if $|\bar{x} - L|$ is largest, γ is given by (2).

Before developing the distribution function of γ, consideration must be given to the situations which might arise in determining γ. Let $\bar{x}$ take on any of the four positions shown below.



If $\bar{x} = \bar{x}_1$ or $\bar{x}_2$, then Equation (1) should be used to find γ. (If (2) were used, the tolerance interval would touch L but not U.) Likewise, if $\bar{x} = \bar{x}_3$ or $\bar{x}_4$, Equation (2) should be used to compute γ. Upon inspection of the above figure, it is apparent that if $\bar{x} < (L + U)/2$, γ should be computed by (1); if $\bar{x} > (L+U)/2$, γ should be computed by Equation (2).

To obtain the distribution function $D(\gamma)$, the identity

$$D(\gamma) = D(\gamma|\bar{x} > (L+U)/2)\Pr(\bar{x} > (L+U)/2)$$
$$+ D(\gamma|\bar{x} < (L+U)/2)\Pr(\bar{x} < (L+U)/2)$$

will be used since the formula for γ is conditional on $\bar{x}$. For notational purposes, let

$$w = \frac{\sqrt{IC\ (1 - \beta)}}{n}$$

$$A \text{ denote } \bar{x} > (L+U)/2$$

$$B \text{ denote } \bar{x} < (L+U)/2.$$

Then

$$D(\gamma|A) = \Pr(\Delta[(w(\bar{x} - L)/s)] \leq \gamma|A)$$
$$= \Pr(w(\bar{x} - L)/s \leq \Delta^{-1}(\gamma)|A).$$

But $(\bar{x} - L)/s$ is not a noncentral-t since the expression is conditional on A (i.e., $\bar{x}$). Hence, let

$$D(\gamma|A) = \Pr(s \geq w(\bar{x} - L)/\Delta^{-1}(\gamma)|A)$$

or

$$D(\gamma|A) = \Pr((n-1)s^2/\sigma^2 \geq (n-1)(w(\bar{x} - L)/\Delta^{-1}(\gamma))^2/\sigma^2|A).$$

From Meyer (21), it is known that if Y has a chi-square distribution with n d.f. that

$$\sqrt{2Y} \stackrel{\cdot}{\sim} N(\sqrt{2n-1}\ ,\ 1)$$

for large n. If $\sqrt{2(n-1)s^2/\sigma^2}$ is denoted by AN, then

$$D(\gamma|A) = Pr(AN \geq \sqrt{2(n-1)}w(\bar{x} - L)/\Delta^{-1}(\gamma)/\sigma|A).$$

If $H(\bar{x})$ is used to denote $\sqrt{2(n-1)}w(\bar{x} - L)/\Delta^{-1}(\gamma)/\sigma$, then

$$D(\gamma|A) = Pr(AN \geq H(\bar{x}),\ \bar{x} > (L + U)/2)/Pr(\bar{x} > (L + U)/2).$$

Now,

$$Pr(\bar{x} < (L + U)/2) = Pr(\sqrt{n}(\bar{x} - (L + U)/2)/s \leq 0)$$

$$= T_a(0)$$

where

$$a = \sqrt{n}(\mu - (L + U)/2)/\sigma.$$

Then

$$D(\gamma|A) = \int_{\bar{x}=(L+U)/2}^{\infty} dN(\mu,\sigma^2/n) \int_{y=H(\bar{x})-\sqrt{2n-1}}^{\infty} dN(0,1)/(1 - T_a(0))$$

$$= \int_{\bar{x}=(L+U)/2}^{\infty} 1 - CN(H(\bar{x}) - \sqrt{2n-1})dN(\mu,\sigma^2/n)/(1 - T_a(0))$$

where

CN(•) is the C.D.F. of a standard normal random variable.

Thus,

$$D(\gamma,A) = \int_{\bar{x}=(L+U)/2}^{\infty} 1 - CN(H(\bar{x}) - \sqrt{2n-1})dN(\mu,\sigma^2/n).$$

In a similar manner, it can be shown that

$$D(\gamma,B) = \int_{-\infty}^{\bar{x}=(L+U)/2} 1 - CN(J(\bar{x}) - \sqrt{2n-1})dN(\mu,\sigma^2/n)$$

where

$$J(x) = \sqrt{2n-1}\, w(U-\bar{x})/\Delta^{-1}(\gamma)/\sigma.$$

Then

$$D(\gamma) = D(\gamma, A) + D(\gamma, B).$$

Efforts to make the form of $D(\gamma)$ more tractable have been unsuccessful. This may be due in part to the approximations used in this development. It would appear that the Wald-Wolfowitz approximate tolerance limits are not well suited to this type of analysis. Similar problems may arise if this type of analysis is tried on other approximate tolerance limits.

CHAPTER VII

SIGNIFICANCE TESTING OF QUANTILE VALUES

In previous chapters, L(S) was defined to be a γ-probability lower OSTL on at least β of the population with density $f(\cdot)$ if

$$\gamma = \Pr\left(\int_{L(S)}^{\infty} f(y)dy \geq \beta\right)$$

$$= \Pr(1 - F(L(S)) \geq \beta)$$

$$= \Pr(F(L(S)) \leq 1 - \beta)$$

$$= \Pr(L(S) \leq F^{-1}(1 - \beta)).$$

Denoting $F^{-1}(1 - \beta)$ by Q_β (i.e., the $1 - \beta$ quantile),

$$\gamma = \Pr(L(S) < Q_\beta). \tag{1}$$

Essentially, Equation (1) states that for a given γ and S, L(S) is a $100\gamma\%$ lower confidence limit on Q_β. According to Bain (13), the value of γ such that L(S) = B, a specified QL, would be the significance level (SL) for testing

$$H: \quad Q_\beta = B$$

versus

$$A: \quad Q_\beta < B.$$

Since

$$F(Q_\beta) = \beta$$

and

$$F(B) = R, \text{ the fraction above } B,$$

the hypothesis H and the alternative A can be restated as

$$H: \quad R = \beta$$

versus

$$A: \quad R < \beta.$$

Upon reflection, it is apparent that the observed confidence values in the previous chapters are the SL values for the significance testing of quantile values.

In the prior chapters, γ was basically of the form

$$\gamma = F(h(S)) \tag{2a}$$

or

$$= \Pr(W \leq h(S)) \tag{2b}$$

where $F(\cdot)$ is the distribution function of a random variable W.

For the normal population with unknown parameters in Chapter III,

$$\gamma = T_\delta(\sqrt{n}(\bar{x} - B)/s)$$

$$= \Pr(t_\delta \leq \sqrt{n}(\bar{x} - B)/s).$$

In the context of significance testing, the test statistic is $\sqrt{n}(\bar{x} - B)/s$ with a null distribution of a noncentral-t with parameter δ and n-1 d.f. Referring back to Chapter III, if $R = \beta$, then the parameter $\theta = \delta$, which verifies that $\sqrt{n}(\bar{x} - B)/s$ has a null distribution with parameter δ. With reference to (2a), it would appear that h(S) would serve as the test statistic with a null distribution function of $F(\cdot)$.

As further verification that γ is actually a SL for quantile testing, consider the form of $d(\gamma)$ when $R = \beta$. Referring again to the normal case in Chapter III with unknown parameters,

$$d(\gamma) = \frac{t_\theta(x)}{t_\delta(x)}\bigg|_{x = T_\delta^{-1}(\gamma)} , \quad 0 < \gamma < 1.$$

If $R = \beta$, then $\theta = \delta$, as stated before, in which case

$$d(\gamma) = 1 , \quad 0 < \gamma < 1.$$

The fact that the SL has a uniform distribution on (0,1) under the null hypothesis is well known (e.g., see (16)). If one were to examine every expression for $d(\gamma)$ in the previous chapters, one would find that $d(\gamma) = 1$ in every case where $R = \beta$.

Distributions of Some Common Significance Levels

It is interesting to note in the Normal case of Chapter III that if $\beta = .5$, then $Q_{.5} = \mu$. If $B = \mu_o$, then the above quantile testing problem becomes a test of

$$H: \quad \mu = \mu_o$$

versus

$$A: \quad \mu < \mu_o$$

with a test statistic $\sqrt{n}(\bar{x} - \mu_o)/s$ and a SL computed by

$$\gamma = T_o(\sqrt{n}(\bar{x} - \mu_o)/s)$$

$$= \Pr(t_o \leq \sqrt{n}(\bar{x} - \mu_o)/s).$$

This is easily recognizable as the SL for the t-test of the mean of a normal distribution. Plots of $d(\gamma)$ for the case $\beta = .05$ are given in Figures 17 and 18. The formula for $d(\gamma)$ is of the form

$$d(\gamma) = \frac{t_\theta(x)}{t_\delta(x)}\bigg|_{x = T_o^{-1}(\gamma)}$$

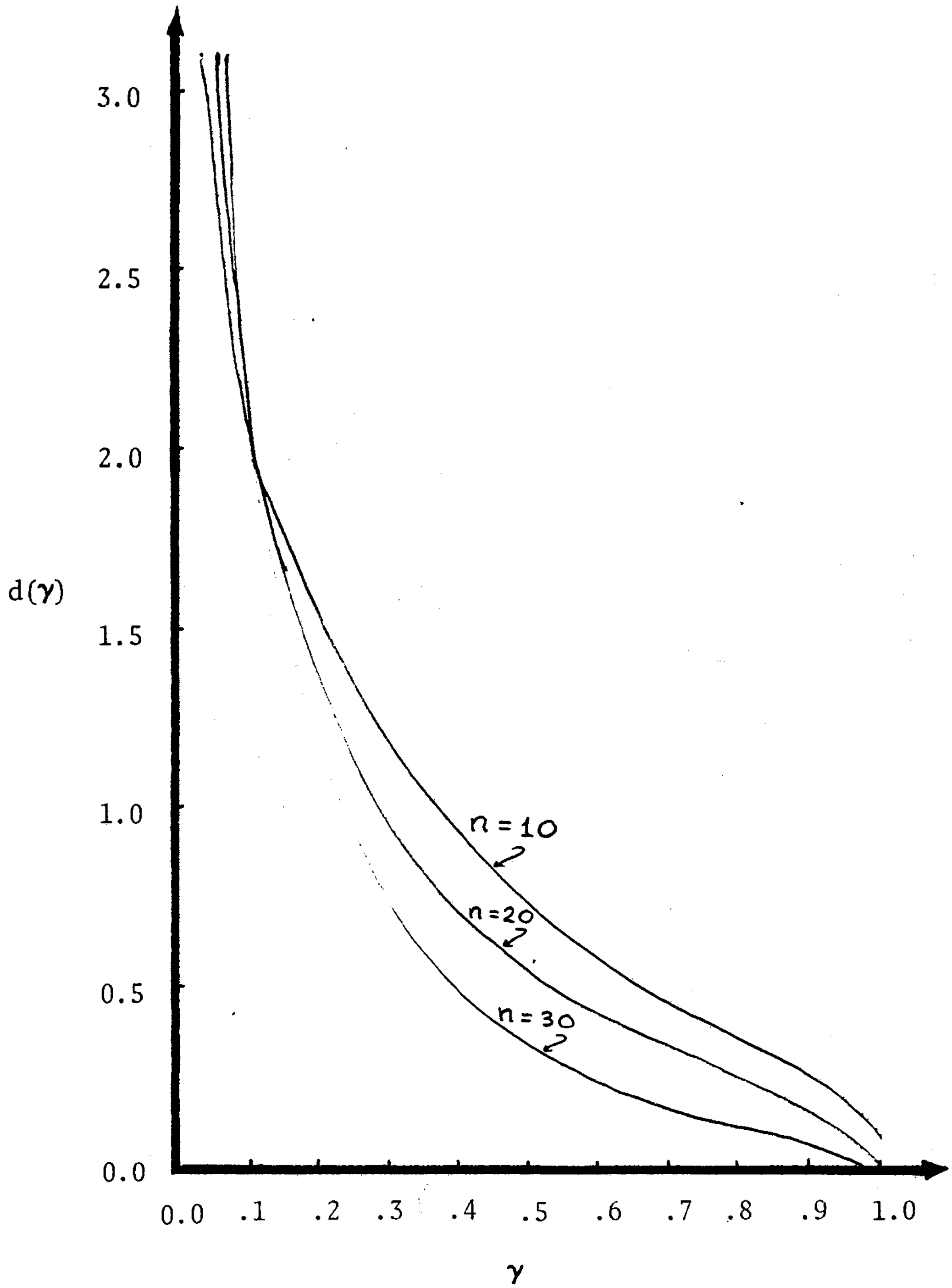


Figure 17. Densities of γ; One-tail t-test; R = .40

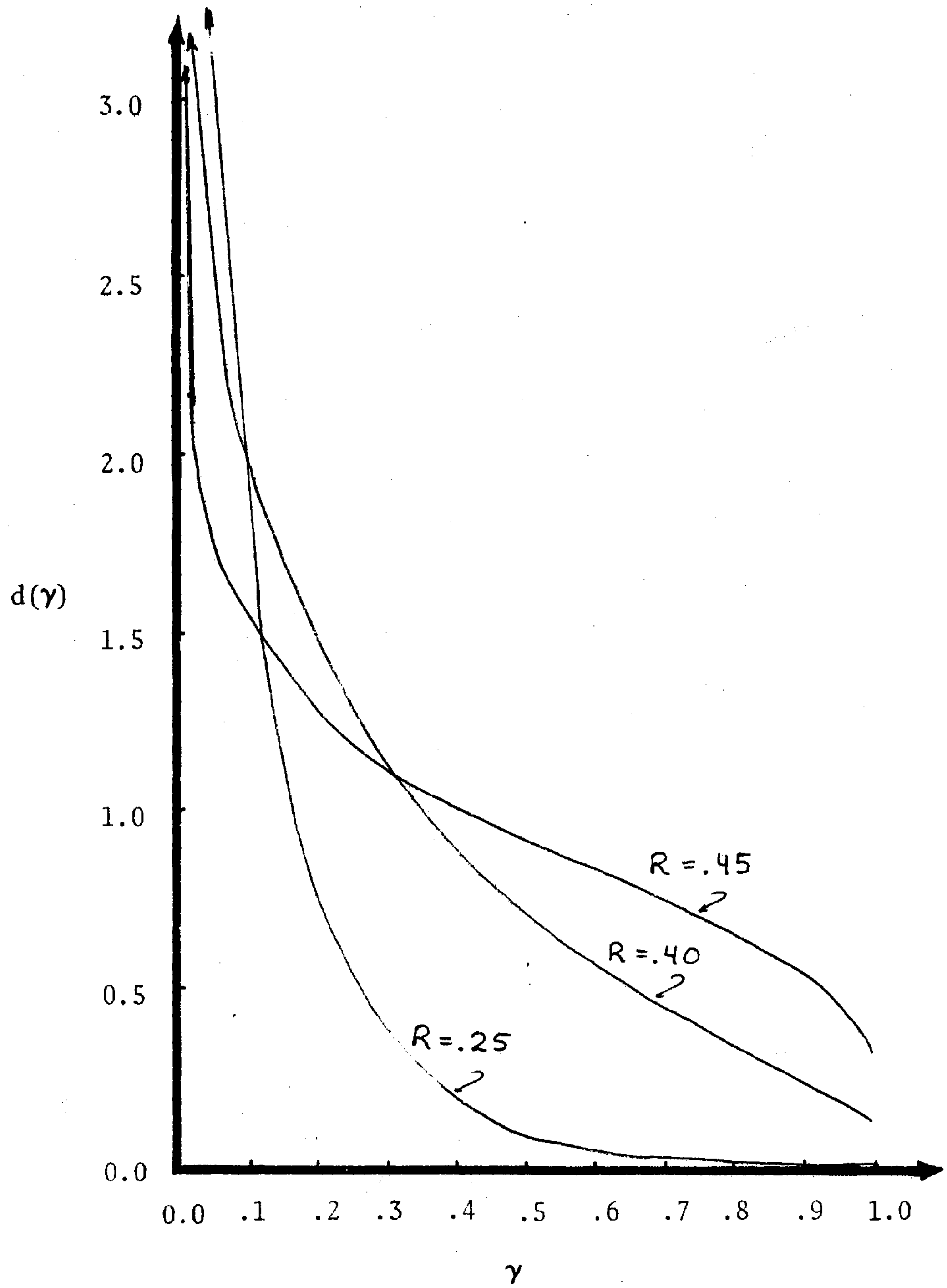


Figure 18. Densities of γ; One-tail t-test; $n = 10$

where

$$\theta = \sqrt{n}\ \text{IN}(R).$$

Likewise, in the normal case where σ is known, letting $\beta = 0.5$ leads to the same set of H and A with the SL given by the expression

$$\gamma = CN_o(\sqrt{n}(\bar{x} - \mu_o)/\sigma)$$

$$= \Pr(Z \leq \sqrt{n}(\bar{x} - \mu_o)/\sigma).$$

The SL has the density

$$d(\gamma) = \frac{f(\text{IN}(\gamma) - \sqrt{n}\ \text{IN}(R))}{f(\text{IN}(\gamma))}$$

where $f(\bullet)$ is the standard normal density. This can be easily reduced to the form

$$d(\gamma) = \exp(\sqrt{n}\ \text{IN}(R)\text{IN}(\gamma) - n\ \text{IN}(R)^2/2)\ ,\ 0 < \gamma < 1.$$

Plots of $d(\gamma)$ for this case are given in Figures 19 and 20.

To compare the curves for σ known and unknown, composite graphs of the $d(\gamma)$ plots were made in Figure 21. Comparison of the curves show remarkably little difference on this scale for even $n = 10$. This would suggest that for testing the mean, the Z-test and t-test are about equally sensitive, unlike the cases where the desired reliability was nearer 1.00, as in Figures 10 and 11. This is a very surprising result and should be more thoroughly pursued.

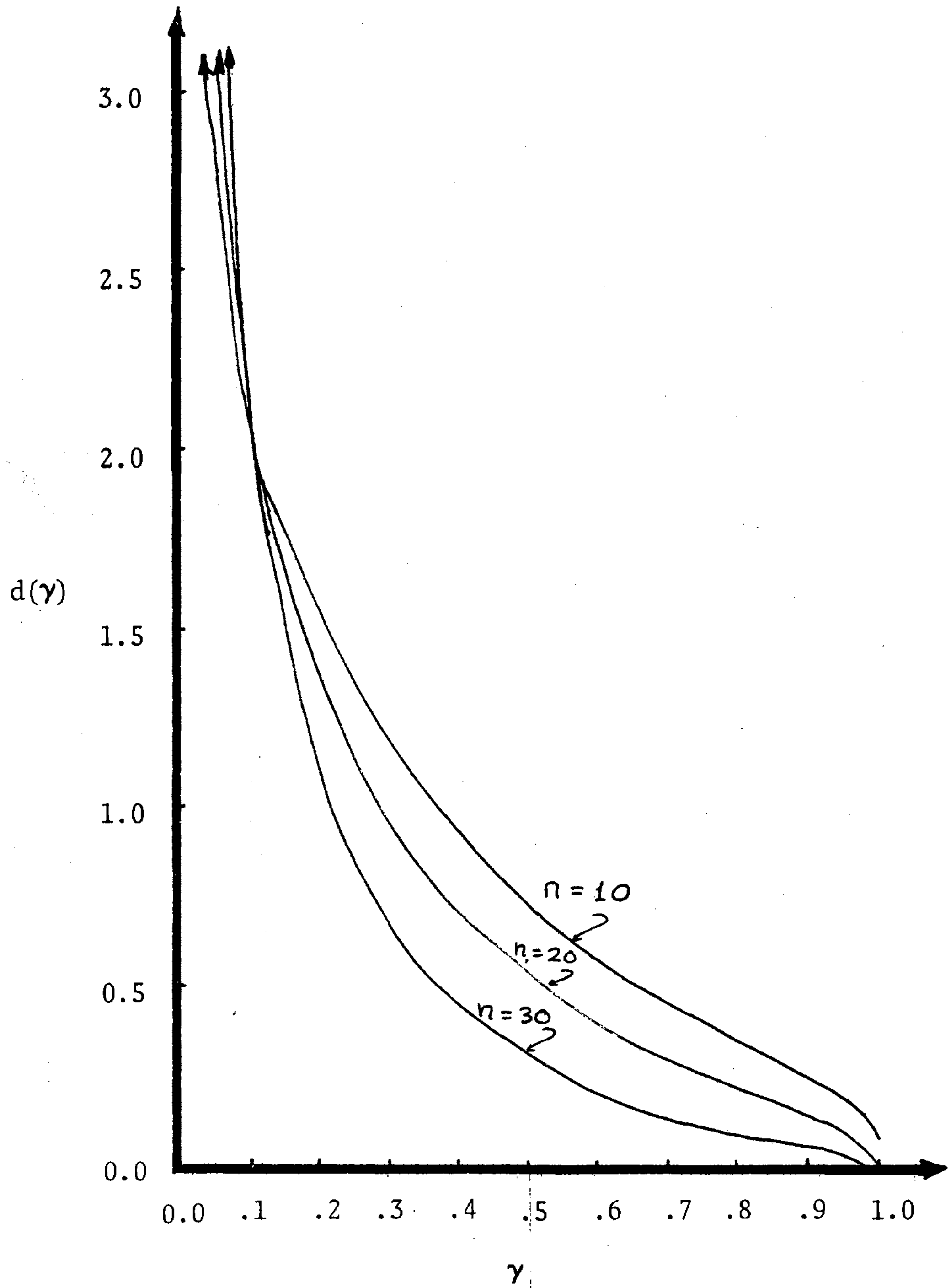


Figure 19. Densities of γ; One-tail Z-test; R = .40

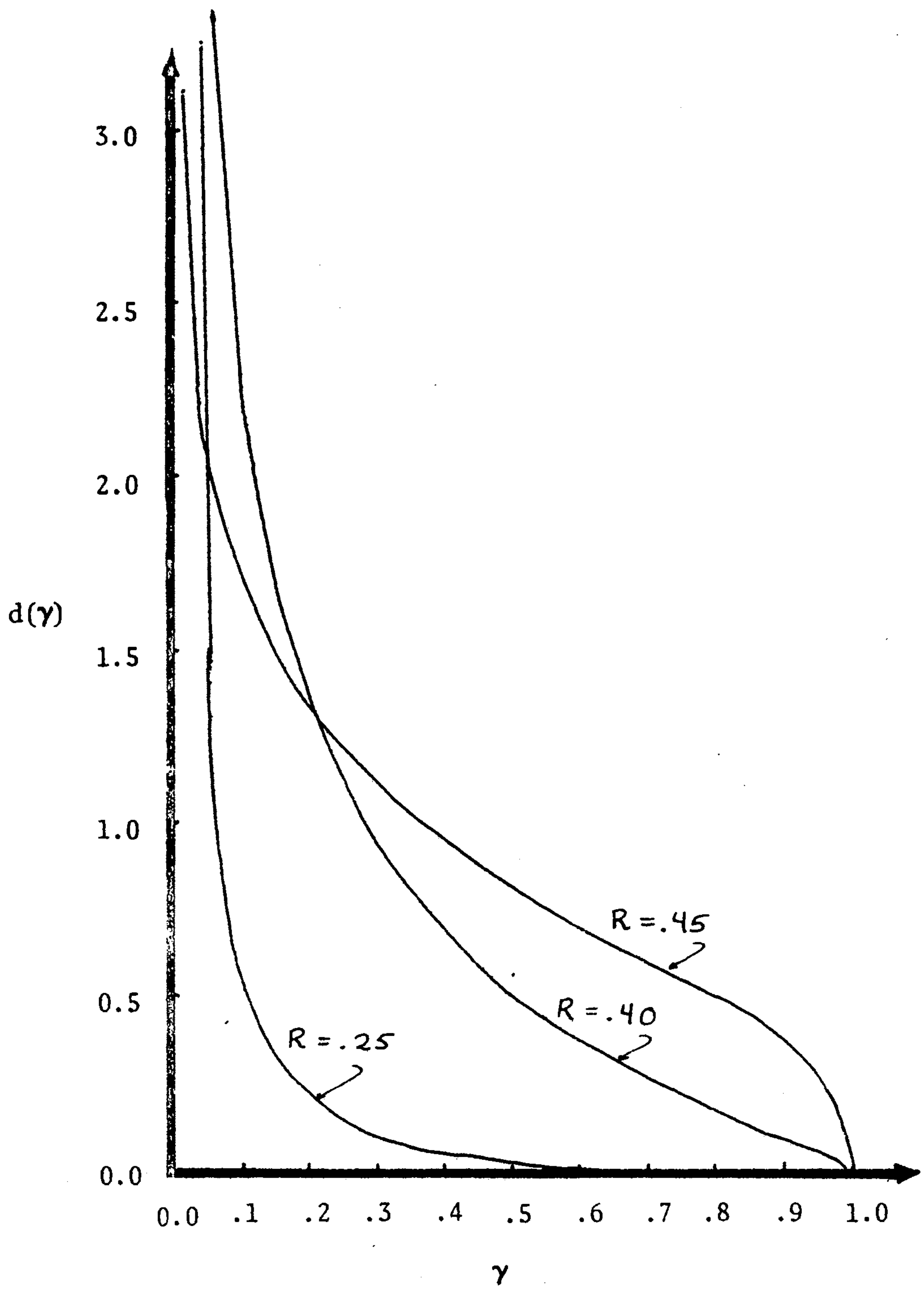


Figure 20. Densities of γ; One-tail Z-test; n = 10

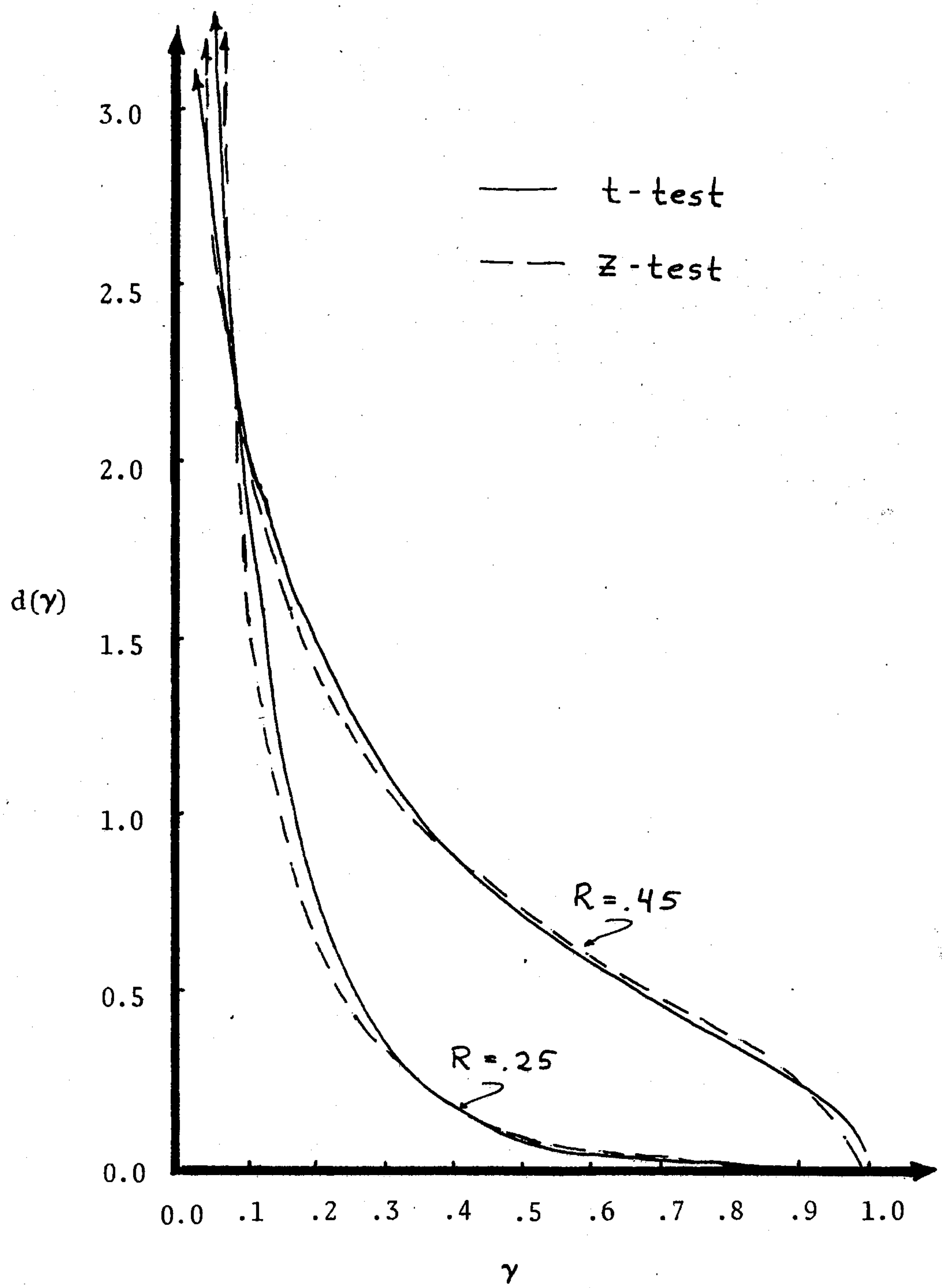


Figure 21. Densities of Y; One-tail t-test and Z-test; n = 10

CHAPTER VIII

EXTENSIONS

The results of the previous chapters were all dependent on the choice of the tolerance limits for each distribution. Certainly other tolerance interval formulas should be used in further investigations to determine their distributional characteristics and compared to the results in this dissertation. Tolerance limits based on order statistics are contained in Bain and Weeks (13).

It would also be worthwhile to investigate the behavior of γ when the assumed distribution for the population is not the correct one. Such errors might be quite common since the normal, gamma, and Weibull densities are not dissimilar in appearance. Attempts to investigate this problem have been unsuccessful due to difficulties with the distribution of the test statistic under a different distribution than that assumed by the statistic. Monte Carlo techniques may be required to investigate this problem.

This writer has been unable to find work similar to that of Chapter VII on the distributional characteristics of the SL. While some work of a more general nature has been done (16), specific details for common significance tests have not been worked out. In addition to expanding the work done here on quantile testing, investigations should also be made into the SL for the chi-square test and F-test for variances.

Throughout this work, the density of γ for continuous parent populations has been generally of the form

$$d(\gamma) = \frac{f_\theta(F_\delta^{-1}(\gamma))}{f_\delta(F_\delta^{-1}(\gamma))} , \quad 0 < \gamma < 1 . \tag{1}$$

It would be interesting to know if this is the general form for the density of γ in all significance tests based on continuous distributions. Furthermore, is $d(\gamma)$ a monotone function for all γ in all significance tests? If it is not, when is it monotone and for what philosophical reason?

The fact that the ratio of two densities in (1) resembles a likelihood ratio suggests an analogous development of significance testing to the theory of tests of hypotheses using the likelihood ratio approach. While the resemblance may only be illusory, a further development in this area might yield more connecting elements between the two theories of testing.

CHAPTER IX

SUMMARY

In reliability assessment today, the engineer will often compute the confidence γ with which he can say that a product is at least $100\beta\%$ reliable, using data from a random sample of products. Several common errors are commonly made in interpreting this observed confidence. First, the engineer often assumes that repeated samples of n products will yield the same observed confidence. In addition, he will interpret the result by saying he will take a $100(1-\gamma)\%$ risk of being wrong if he says that the product is not $100\beta\%$ reliable (with respect to the preset specifications).

This thesis has shown that γ is actually an observed significance level for determining how well the data conform to the notion that the product is at least $100\beta\%$ reliable, as opposed to being less than $100\beta\%$ reliable. If the observed confidence is near 1.00, he should consider this to be strong evidence that the product is at least $100\beta\%$ reliable; conversely, if the observed confidence is near 0.0, he should consider this as casting doubt on the product being at least $100\beta\%$ reliable.

To find the observed confidence γ that a product is at least $100\beta\%$ reliable, tolerance limits are used in conjunction with the specification or qualification limit for the product. The observed confidence would be the value of γ which would be quired to have the tolerance limit equal the qualification limit. Both one-sided and two-sided tolerance limits were used to find the observed confidence, with the most meaningful results coming from the use of one-sided tolerance limits (OSTL's).

The lower OSTL's investigated were those based on the normal, gamma, Weibull, and exponential distributions. Additional work was done on lower tolerance limits for binomial and Poisson distributions. In all cases, the results showed that if the true reliability (R) exceeded the desired reliability (β), values of γ near 1.00 were the most likely; conversely, if the true reliability was less than the desired reliability, values of γ near 0.00 were the most likely. The skewness of the density function of γ increased as the sample size increased and/or as the disparity between the true reliability and desired reliability increased.

The results for the normal with known and unknown variance indicated that a knowledge of the variance led to a more sensitive test statistic in γ. That is, γ was more likely to be near 1.00 when $R > \beta$ when σ^2 was known than when σ^2 was unknown. An attempt to appeal to the Central Limit Theorem and use the normal tolerance limits on Weibull data showed a significant difference in results and possible conclusions. These results suggest that failure to use any available information about the population from which the sample was drawn may result in a marked difference in findings from the data.

Investigations were also made into the use of γ as an observed significance level (SL) for significance testing of quantile values. If the desired reliability was set equal to 0.5, the value of γ in the normal cases became a SL for one-tail tests of the mean. Preliminary results show little difference in the sensitivity of the one-tail t-test and Z-test on the normal mean. This is an unexpected result that needs further verification before being accepted.

A SELECTED BIBLIOGRAPHY

(1) Mosteller, Fredrick. "Note on an Application of Runs to Quality Control Charts." Annals Math. Stat., Vol. 12 (1941), 228.

(2) Neyman, J. "On a Statistical Problem Arising in Routine Analyses and in Sampling Inspection of Mass Production." Annals Math. Stat, Vol. 12 (1941), 46-68.

(3) Jennett, W. J., and B. L. Welch. "The Control of Proportion Defective as Judged by a Single Quality Characteristic Varying on a Continuous Scale." Jour. Royal Stat. Soc., Vol. 6 (1939), 80-95.

(4) Dodge, H. F. "A Sampling Inspection Plan for Continuous Production." Annals Math Stat., Vol. 14 (1943), 264.

(5) Wolfowitz, J. "On a Theory of Runs With Some Application to Quality Control." Annals Math. Stat., Vol. 14 (1943), 280.

(6) Wald, A., and J. Wolfowitz. "Sampling Inspection Plans for Continuous Production Which Insure a Prescribed Limit on the Outgoing Quality." Annals Math. Stat., Vol. 16 (1945), 30.

(7) Wald, A., and J. Wolfowitz. "Tolerance Limits for a Normal Distribution." Annals Math. Stat., Vol. 17 (1946), 208.

(8) Bowker, A. H. "Computation of Factors for Tolerance Limits on a Normal Distribution When the Sample is Large." Annals Math. Stat., Vol. 17 (1946), 238.

(9) Wilks, S. S. "Determination of Sample Sizes for Setting Tolerance Limits." Annals Math. Stat. (1941), 91.

(10) Wald, A. "Setting of Tolerance Limits When the Sample is Large." Annals Math. Stat., Vol. 13 (1942), 208.

(11) Scheffe, H., and J. W. Tukey. "A Formula for Sample Sizes for Population Tolerance Limits." Annals Math. Stat., Vol. 15 (1944), 217.

(12) Goodman, L. A., and A. Madansky. "Parameter-Free and Nonparametric Tolerance Limits: The Exponential Case." Technometrics, Vol. 4 (1962), 75-95.

(13) Bain, Lee J., and D. L. Weeks. "Tolerance Limits for the Generalized Gamma Distribution." Jour. Amer. Stat. Assoc., Vol. 60 (1965), 1142-1152.

(14) Mendenhall, W. "A Bibliography on Life Testing and Related Topics." Biometrika, Vol. 45 (1958), 521-543.

(15) Martin Marietta Corporation. Reliability for the Engineer, Book Five: Testing for Reliability, 1966, 11.

(16) Kempthorne, O., and Leroy Folks. Probability, Statistics, and Data Analysis. Iowa State Press, 1971.

(17) Anscombe, F. J. "Tests of Goodness of Fit." Jour. Royal Stat. Soc., Series B, Vol. 25 (1963), 81-94.

(18) Fisher, R. A. Statistical Methods and Scientific Inference. New York: Hafner Publishing Co., 1956.

(19) Owen, D. B. "A Survey of Properties and Applications of the Noncentral t-Distribution." Technometrics, Vol. 10 (1968), 295.

(20) Resnikoff, G., and G. Liebermann. Tables of the Non-Central T Distribution. Stanford University Press, 1959.

(21) Meyer, Paul L. Introductory Probability and Statistical Applications. Reading, Mass.: Addison-Wesley Publishing Company, 1965.

APPENDIX

Given $f(x)$ to be an integrable function of x with $f(-\infty) = 0$. Let $F^{-1}(\gamma)$ be defined by

$$\gamma = \int_{-\infty}^{F^{-1}(\gamma)} f(x)\,dx \tag{1}$$

for all γ between zero and one.

Differentiating (1) with respect to γ,

$$1 = f(F^{-1}(\gamma))\; dF^{-1}(\gamma)/d\gamma$$

or

$$dF^{-1}(\gamma)/d\gamma = 1/f(F^{-1}(\gamma)).$$

VITA

Richard Harold Browne

Candidate for the Degree of

Doctor of Philosophy

Thesis: DISTRIBUTIONAL CHARACTERISTICS OF THE OBSERVED CONFIDENCE COEFFICIENT IN CERTAIN RELIABILITY PROBLEMS

Major Field: Statistics

Biographical:

Personal Data: Born in Clayton, Missouri, September 24, 1946, the son of Basil and Evelyn Browne.

Education: Attended elementary school in Brentwood and Kirkwood, Missouri; attended high school in Kirkwood, Missouri; graduated from Kirkwood Senior High School in 1964; received the Associate of Science degree from Meramec Community College in June, 1966; received the Bachelor of Science degree from the University of Missouri at Rolla with a major in Applied Mathematics in May, 1968; received the Master of Science degree from Oklahoma State University with a major in Statistics in May, 1970; completed requirements for the Doctor of Philosophy degree in May, 1973 at Oklahoma State University.

Professional Experience: Employed as computer programmer by Nooter Corporation, St. Louis, Missouri, from June, 1967 to September, 1969; graduate assistant at Oklahoma State University from September, 1968 to May, 1973; statistician at M. D. Anderson Hospital and Tumor Institute, Houston, Texas from June, 1972 to September, 1972.

Professional Organizations: American Statistical Association, Biometric Society.